DOCUMENT RESUME

ED 376 191 TM 022 279

AUTHOR	Johnson, Eugene G.; And Others
TITLE	The NAEP 1992 Technical Report.
INSTITUTION	Educational Testing Service, Princeton, N.J.; National Assessment of Educational Progress, Princeton, NJ.
SPONS AGENCY	National Center for Education Statistics (ED), Washington, DC.
REPORT NO	ISBN-0-88685-161-0; NAEP-R-23-TR20; NCES-94-490
PUB DATE	Jul 94
NOTE	1,131p.
PUB TYPE	Reports - Evaluative/Feasibility (142) -- Statistical Data (110)
EDRS PRICE	MF08/PC46 Plus Postage.
DESCRIPTORS	Academic Achievement; Data Analysis; *Data Collection; *Educational Assessment; Elementary Secondary Education; Grade 4; Grade 8; Grade 11; Grade 12; Mathematics; National Surveys; Private Schools; Public Schools; Research Design; *Research Methodology; *Sampling; Scaling; Sciences; Scoring; Statistical Studies; *Test Construction; Test Items; Writing (Composition)
IDENTIFIERS	*National Assessment of Educational Progress; Weighting (Statistical)

ABSTRACT

The 1992 National Assessment of Educational Progress (NAEP) monitored the performance of students in American schools in reading, mathematics, science, and writing. The sample consisted of more than 145,000 public and private school students in grades 4, 8, 11, and 12. This technical report provides details of instrument development, sample design, and data collection and analysis. The following chapters describe technical aspects of the survey: (1) "Overview of Part I: The Design and Implementation of the 1992 NAEP" (Eugene G. Johnson); (2) "Developing the NAEP Objectives, Items, and Background Questions for the 1992 Assessments of Reading, Mathematics, and Writing" (Jay R. Campbell, and others); (3) "Sample Design" (Keith F. Rust and Leslie Wallace); (4) "Assessment Instruments" (Debra L. Kline, and others); (5) "Field Operations and Data Collection" (Nancy Caldwell and Lesly Fleming); (6) "Processing of Assessment Materials" (Dianne Smrdel, and others); (7) "Professional Scoring" (Andrew S. Latham); (8) "Database Creation, Quality Control of Data Entry, and Database Products" (Alfred M. Rogers, and others); (9) "Overview of Part II: The Analysis of 1992 NAEP Data" (James E. Carlson and Spencer S. Swinton); (10) "Weighting Procedures and Estimation of Sampling Variance" (Eugene G. Johnson, and others); (11) "Scaling Procedures" (Eugene G. Johnson, and others); (12) "Data Analysis for the Reading Assessment" (John R. Donoghue, and others); (13) "Data Analysis for the Mathematics Assessment" (Frank Jenkins and Edward M. Kulick); (14) "Data Analysis for the Science Long-term Trend Assessment" (Nancy L. Allen and Steven P. Isham); (15) "Data Analysis for the Writing Assessment" (James E. Carlson, and others); and (16) "Statistical Summary of 1992 NAEP Data and Estimates of the Proficiencies of American Students" (Bruce A. Kaplan, and others). Statistical data are presented in 306 tables and 24 figures. Sixteen appendixes provide specific additional statistical background. (Contains 117 references.) (SLD)

ED 376 191

NATIONAL CENTER FOR EDUCATION STATISTICS

The NAEP
1992 Technical Report

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

JOURNAL
HOW TO...
POETRY
SHORT STORIES

TM022279

THE NATION'S REPORT CARD naep

Prepared by Educational Testing Service under contract with the National Center for Education Statistics.

Office of Educational Research and Improvement
U.S. Department of Education

BEST COPY AVAILABLE

What is The Nation's Report Card?

THE NATION'S REPORT CARD, the National Assessment of Educational Progress (NAEP), is the only nationally representative and continuing assessment of what America's students know and can do in various subject areas. Since 1969, assessments have been conducted periodically in reading, mathematics, science, writing, history/geography, and other fields. By making objective information on student performance available to policymakers at the national, state, and local levels, NAEP is an integral part of our nation's evaluation of the condition and progress of education. Only information related to academic achievement is collected under this program. NAEP guarantees the privacy of individual students and their families.

NAEP is a congressionally mandated project of the National Center for Education Statistics, the U.S. Department of Education. The Commissioner of Education Statistics is responsible, by law, for carrying out the NAEP project through competitive awards to qualified organizations. NAEP reports directly to the Commissioner, who is also responsible for providing continuing reviews, including validation studies and solicitation of public comment, on NAEP's conduct and usefulness.

In 1988, Congress created the National Assessment Governing Board (NAGB) to formulate policy guidelines for NAEP. The board is responsible for selecting the subject areas to be assessed, which may include adding to those specified by Congress; identifying appropriate achievement goals for each age and grade; developing assessment objectives; developing test specifications; designing the assessment methodology; developing guidelines and standards for data analysis and for reporting and disseminating results; developing standards and procedures for interstate, regional, and national comparisons; improving the form and use of the National Assessment; and ensuring that all items selected for use in the National Assessment are free from racial, cultural, gender, or regional bias.

The National Assessment Governing Board

Mark D. Musick, Chairman
President
Southern Regional Education Board
Atlanta, Georgia

Hon. William T. Randall, Vice Chair
Commissioner of Education
State Department of Education
Denver, Colorado

Parris C. Battle
Education Specialist
Dade County Public Schools
Miami, Florida

Honorable Evan Bayh
Governor of Indiana
Indianapolis, Indiana

Mary R. Blanton
Attorney
Blanton & Blanton
Salisbury, North Carolina

Linda R. Bryant
Dean of Students
Florence Reizenstein Middle School
Pittsburgh, Pennsylvania

Naomi K. Cohen
Office of Policy and Management
State of Connecticut
Hartford, Connecticut

Charlotte Crabtree
Professor
University of California
Los Angeles, California

Chester E. Finn, Jr.
Founding Partner and Senior Scholar
The Edison Project
Washington, DC

Michael J. Guerra
Executive Director
NCEA, Secondary School Department
Washington, District of Columbia

William (Jerry) Hume
Chairman of the Board
Basic American, Inc.
San Francisco, California

Christine Johnson
Director of Urban Initiatives
Education Commission of the States
Denver, Colorado

John S. Lindley
Principal
Galloway Elementary School
Henderson, Nevada

Jan B. Loveless
Jan B. Loveless and Associates
Educational Consultants
Midland, Michigan

Marilyn McConachie
Member, Board of Education
Glenbrook High Schools
Northbrook, Illinois

Honorable Stephen E. Merrill
Governor of New Hampshire
Concord, New Hampshire

Jason Millman
Professor
Cornell University
Ithaca, New York

Honorable Richard P. Mills
Commissioner of Education
State Department of Education
Montpelier, Vermont

John A. Murphy
Superintendent of Schools
Charlotte-Mecklenburg Schools
Charlotte, North Carolina

Mitsugi Nakashima
Hawaii State Board of Education
Honolulu, Hawaii

Michael T. Nettles
Professor
University of Michigan
Ann Arbor, Michigan

Honorable Edgar D. Ross
Senator
Legislature of the Virgin Islands
Christiansted, St. Croix
U.S. Virgin Islands

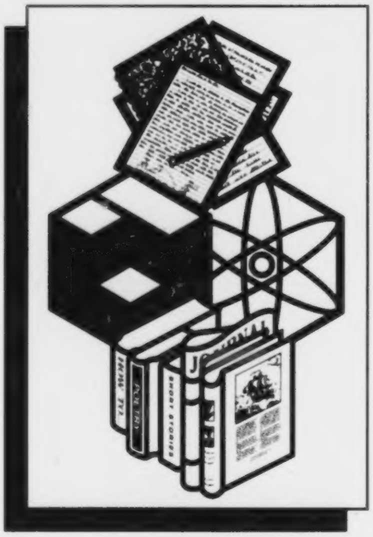
Marilyn Whirry
English Teacher
Mira Costa High School
Manhattan Beach, California

Sharon P. Robinson
Assistant Secretary for Educational
Research and Improvement (Ex-Officio)
U.S. Department of Education
Washington, D.C.

Roy Truby
Executive Director, NAGB
Washington, D.C.

NATIONAL CENTER FOR EDUCATION STATISTICS

The NAEP 1992 Technical Report



Eugene G. Johnson and James E. Carlson
in collaboration with

Nancy L. Allen
Mary Lyn Bourque
Drew W. Bowker
Nancy W. Caldwell
Jay R. Campbell
John R. Donoghue
John J. Ferris
Lesly Flemming
Yim Fai Fong
David S. Freund
Steven P. Isham
Frank Jenkins
Laura S. Jerry
Bruce A. Kaplan
Debra L. Kline
Edward M. Kulick
Andrew S. Latham
Stephen Lazer
Phillip T. Leung
John Mazzeo
Robert J. Mislevy
Lavonne Mohn
Ina V. S. Mullis
Eiji Muraki
Robert Neiers
Linda Reynolds
Alfred M. Rogers
Keith F. Rust
Connie Smith-Hora
Dianne Smrdel
Brent Studer
Spencer S. Swinton
Brad Thayer
Neal Thomas
Leslie Wallace

Report No. 23-TR20 July 1994

THE NATION'S REPORT CARD naep

Prepared by Educational Testing Service under contract
with the National Center for Education Statistics

Office of Educational Research and Improvement
U.S. Department of Education

U.S. Department of Education
Richard W. Riley
Secretary

Office of Educational Research and Improvement
Sharon P. Robinson
Assistant Secretary

National Center for Education Statistics
Emerson J. Elliott
Commissioner

FOR MORE INFORMATION:

For ordering information on this report, write:

Education Information Branch
Office of Educational Research and Improvement
U.S. Department of Education
555 New Jersey Avenue, NW
Washington, D.C. 20208-5641

or call 1-800-424-1616 (in the Washington, D.C. metropolitan area call 202-219-1651).

ISBN: 0-88685-161-0

The work upon which this publication is based was performed for the National Center for Education Statistics, Office of Educational Research and Improvement, by Educational Testing Service.

Educational Testing Service is an equal opportunity, affirmative action employer.

Educational Testing Service, *ETS*, and *the ETS logo* are registered trademarks of Educational Testing Service.

TM

THE NAEP 1992 TECHNICAL REPORT

CONTENTS

List of Tables xi
List of Figures xxi
Acknowledgments xxiii

Introduction 3
Eugene G. Johnson and James E. Carlson

PART I
The Design and Implementation of the 1992 NAEP

Chapter 1 Overview of Part I: The Design and Implementation of the 1992 NAEP 9
Eugene G. Johnson

- 1.1 The 1992 NAEP Design 10
  - 1.1.1 The 1992 NAEP Samples 13
  - 1.1.2 NAEP Assessments Since 1969 15
- 1.2 Development of Assessment Objectives, Items, and Background Questions 20
- 1.3 The 1992 Sample Design 20
- 1.4 Assessment Instruments 23
  - 1.4.1 Student Assessment Booklets—Main Assessment 23
  - 1.4.2 Student Assessment Booklets—Long-term Trend Samples 29
  - 1.4.3 Other Instruments 29
- 1.5 Field Operations and Data Collection 30
- 1.6 Materials and Data Processing 31
- 1.7 Professional Scoring 31
- 1.8 Creation of the Database 31

Chapter 2 Developing the NAEP Objectives, Items, and Background Questions for the 1992 Assessments of Reading, Mathematics, and Writing 33
Jay R. Campbell, Stephen Lazer, and Ina V.S. Mullis

- 2.1 General Overview of the 1992 Assessment Frameworks 35
- 2.2 General Overview of Procedures for Developing the Items 36
- 2.3 Developing the Reading Assessment 37
  - 2.3.1 Overview 38
  - 2.3.2 Development of the Assessment Framework 38
  - 2.3.3 Framework and Assessment Design Principles 39
  - 2.3.4 Framework for the 1992 Assessment 40
  - 2.3.5 Developing the Cognitive Items 43
  - 2.3.6 Distribution of Assessment Items 44

2.3.7 Reading Background Questionnaires 46
2.3.8 Development of Final Forms 48
2.4 Developing the Mathematics Assessment 48
2.4.1 Overview 48
2.4.2 Development of the Assessment Framework 49
2.4.3 Framework and Assessment Design Principles 50
2.4.4 Framework for the 1992 Assessment 51
2.4.5 Developing the Cognitive Items 51
2.4.6 Distribution of Assessment Items 55
2.4.7 Mathematics Background Questionnaires 57
2.4.8 Development of Final Forms 58
2.5 Developing the Writing Assessment 59
2.5.1 Overview 59
2.5.2 Development of the Assessment Framework 59
2.5.3 Framework and Assessment Design Principles 60
2.5.4 Framework for the 1992 Assessment 61
2.5.5 Developing the Cognitive Items 61
2.5.6 Writing Background Questionnaires 63
2.5.7 Development of Final Forms 64
2.6 Background Questionnaires 64
2.6.1 Student Questionnaires 65
2.6.2 Teacher, School, and Excluded Student Questionnaires 65

Chapter 3 Sample Design 67
Keith F. Rust and Leslie Wallace

3.1 Primary Sampling Units 69
3.2 Schools for Main Samples; Assignment of Sessions to Schools 71
3.3 Selection of Schools for Long-term Trend Samples; Assignment of Sessions to Schools 77
3.4 Sampling Students 79
3.5 Excluded Students 80
3.6 Student Participation Rates 83
3.7 Overall Student Participation Rates 83
3.8 Sampling Teachers 86

Chapter 4 Assessment Instruments 87
Debra L. Kline, Stephen Lazer, and Jay R. Campbell

4.1 Student Booklets—Main Assessments 87
4.1.1 Reading 87
4.1.2 Mathematics 92
4.1.3 Writing 99
4.2 Student Booklets—Long-term Trend Assessments 100
4.3 Main Assessment Student Questionnaires 100
4.3.1 Reading Background Questionnaire 108
4.3.2 Mathematics Background Questionnaire 108
4.3.3 Writing Background Questionnaire 108

4.4 Teacher Questionnaires 109
- 4.4.1 Reading Teacher Questionnaire 109
- 4.4.2 Mathematics Teacher Questionnaire 109
- 4.4.3 Writing Teacher Questionnaire 110

4.5 Excluded Student Questionnaire 110
4.6 School Characteristics and Policies Questionnaire 110
4.7 Principal Questionnaire 111
4.8 The Writing Portfolio Study 111

Chapter 5 Field Operations and Data Collection 113
Nancy Caldwell and Lesly Flemming

5.1 Schedule of Activities 113
5.2 Training Field Staff 114
- 5.2.1 Supervisor Training 114
- 5.2.2 Recruiting, Hiring, and Training Exercise Administrators 115

5.3 Securing Cooperation in Sampled Schools 116
- 5.3.1 Introductory Meetings 116
- 5.3.2 Making Arrangements for the Assessments 117

5.4 Long-term Trend Assessments 118
- 5.4.1 Overview 118
- 5.4.2 Selecting the Student Sample 119
- 5.4.3 Conduct of the Assessment 120
- 5.4.4 Results of the Long-term Trend Assessments 122
- 5.4.5 Assessment Questionnaires and Reports

5.5 Main Assessments 124
- 5.5.1 Overview 124
- 5.5.2 Selecting the Student Sample 124
- 5.5.3 Integrated Reading Performance Record (IRPR) 127
- 5.5.4 The Writing Portfolio Study 127
- 5.5.5 Results of the Assessment 127
- 5.5.6 Assessment Questionnaires and Reports 128

5.6 Field Management 128

Chapter 6 Processing of Assessment Materials 131
Dianne Smrdel, Linda Reynolds, and Brad Thayer

6.1 Overview 131
6.2 Process Control System 132
6.3 Workflow Management System 132
6.4 Process Flow of Materials and Database Creation 132
6.5 Materials Distribution 134
6.6 Processing Assessment Materials 135
6.7 Data Transcription Systems 137
- 6.7.1 Data Entry 138
- 6.7.2 Scanning 138

6.8 Data Validation 139
6.9 Editing 140
6.10 Questionnaires 142

6.11 Merging of Student Data 142
6.12 Storage of Documents 142

Chapter 7 Professional Scoring 143
Andrew S. Latham

7.1 Main and Trial State Mathematics Assessments 144
  7.1.1 Characteristics of the Scoring Guides 144
  7.1.2 Training 146
  7.1.3 Scoring 146
7.2 Main and Trial State Reading Assessments 147
  7.2.1 Characteristics of the Scoring Guides 147
  7.2.2 Training 148
  7.2.3 NAEP Reader 148
  7.2.4 Integrated Reading Performance Record 149
7.3 Main Writing Assessment 151
  7.3.1 Characteristics of the Scoring Guides 151
  7.3.2 Training 152
  7.3.3 Scoring 152
7.4 Long-term Trend Assessments 152
  7.4.1 Characteristics of the Scoring Guides 152
  7.4.2 Training 153
  7.4.3 Scoring 153
  7.4.4 Holistic Scoring of Writing 155
  7.4.5 Writing Mechanics Scoring 155
  7.4.6 Writing Portfolio Scoring 157

Chapter 8 Database Creation, Quality Control of Data Entry, and Database Products 159
Alfred M. Rogers, David S. Freund, John J. Ferris

8.1 Creation of the Database 159
  8.1.1 Merging Files 159
  8.1.2 Creating the Master Catalog 160
8.2 Quality Control of Data Entry 161
  8.2.1 Student Data 164
  8.2.2 Excluded Student Questionnaire Data 164
  8.2.3 Teacher Questionnaire Data 164
  8.2.4 School Characteristics and Policies Questionnaire Data 165
8.3 NAEP Database Products 165
  8.3.1 The Item Information Database 165
  8.3.2 The Restricted-use Data Files 166
  8.3.3 The Secondary-use Data Files 166

PART II
The Analysis of 1992 NAEP Data

Chapter 9 Overview of Part II: The Analysis of 1992 NAEP Data 175
James E. Carlson and Spencer S. Swinton

9.1 Samples of Students 176
9.2 Analysis Steps 177
  9.2.1 Preparation of Final Sampling Weights 177
  9.2.2 Analysis of Item Properties: Background and Cognitive Items 177
  9.2.3 Scaling 180
  9.2.4 Drawing Inferences from the Results 184
  9.2.5 Minimum Sample Sizes for Reporting Subgroup Results 186
  9.2.6 Estimates of Standard Errors with Large Mean Squared Errors 186
  9.2.7 Teacher Questionnaires 187
  9.2.8 Dimensionality Analysis 187
9.3 Overview of Chapters 10 Through 15 190

Chapter 10 Weighting Procedures and Estimation of Sampling Variance 193
Eugene G. Johnson, Keith F. Rust, and Leslie Wallace

10.1 Derivation of the Sample Weights 196
  10.1.1 Student Base Weights 197
  10.1.2 Adjustment of Base Weights for Nonresponse 197
  10.1.3 Trimming of Weights 211
  10.1.4 Poststratification 213
  10.1.5 The Final Student Weight 215
  10.1.6 Other Weights 216
10.2 Procedures Used by NAEP to Estimate Sampling Variability 218
10.3 Approximating the Sampling Variance Using Design Effects 222
  10.3.1 Design Effects for Proportion-correct Statistics 223
10.4 The Degrees of Freedom of the Variance Estimate 236

Chapter 11 Scaling Procedures 241
Eugene G. Johnson, Robert J. Mislevy, and Neal Thomas

11.1 Overview 241
11.2 Background 241
11.3 Scaling Methodology 243
  11.3.1 The Scaling Models 243
  11.3.2 An Overview of Plausible Values Methodology 247
  11.3.3 Computing Plausible Values in IRT-based scales 249
11.4 Achievement Levels and Scale Anchoring 250
11.5 Analyses 251
  11.5.1 Computational Procedures 252
  11.5.2 Statistical Tests 253

11.5.3 Biases in Secondary Analyses 254
11.5.4 A Numerical Example 255
11.6 Overview of the 1992 NAEP Scales 256

Chapter 12 Data Analysis for the Reading Assessment 257
John R. Donoghue, Steven P. Isham, Drew W. Bowker, and David S. Freund

12.1 Long-term Trend Data Analysis 258
12.1.1 Item Analysis for the Long-term Trend Assessment 262
12.1.2 Treatment of Constructed-response Items 262
12.1.3 Scaling 262
12.1.4 Generation of Plausible Values 266
12.1.5 The Final Long-term Trend Proficiency Scale 269
12.1.6 Partitioning of the Estimation Error Variance 271
12.2 Main Data Analysis 271
12.2.1 Overview 271
12.2.2 Description of Items and Assessment Booklets 272
12.2.3 Item Analyses 273
12.2.4 IRT Scaling 282
12.2.5 Generation of Plausible Values 292
12.2.6 The Final Proficiency Scales 295
12.2.7 Partitioning of the Estimation Error Variance 296

Chapter 13 Data Analysis for the Mathematics Assessment 299
Frank Jenkins and Edward M. Kulick

13.1 Long-term Trend Data Analysis 301
13.1.1 Item Analysis for the Long-term Trend Assessment 304
13.1.2 Item Parameter Estimation 306
13.1.3 Derived Background Variables 306
13.1.4 Generation of Plausible Values 309
13.1.5 The Final Proficiency Scale 309
13.2 Main Data Analysis 310
13.2.1 The Main Focused-BIB Assessment 311
13.2.2 The Calculator Bridge Assessment for Grade 4 313
13.2.3 Special Mathematics Assessment 315
13.2.4 Item Analysis 318
13.2.5 Constructed-response Items 322
13.2.6 Differential Item Function Analysis 323
13.2.7 Estimation of Item Parameters 324
13.2.8 Evaluating the Fit of the IRT Model 327
13.2.9 Derived Background Variables 327
13.2.10 Generation of Plausible Values 331
13.2.11 The Transformation of the Proficiency Scale 332
13.2.12 Analysis of Short-term Trend 334
13.2.13 Partitioning of the Estimation Error Variance 337
13.2.14 Mathematics Teacher Questionnaire 337
13.2.15 Analysis of Dimensionality 337

13.2.16 Anchoring the Points on the Mathematics Proficiency Scale 340
13.2.17 Grade 4: Comparing the Four-scale Composite of 1990 with the Five-scale Composite of 1992 341

Chapter 14 Data Analysis for the Science Long-term Trend Assessment 343
Nancy L. Allen and Steven P. Isham

14.1 Item Analysis 347
14.2 Estimation of Item Parameters 347
14.3 Derived Background Variables 350
14.4 Generation of Plausible Values 350
14.5 The Final Proficiency Scale 352
14.6 Partitioning of the Estimated Error Variance 352
14.7 Anchoring the Points on the Science Proficiency Scale 353

Chapter 15 Data Analysis for the Writing Assessment 355
James E. Carlson, Eiji Muraki, Bruce A. Kaplan, and Yim Fai Fong

15.1 Long-term Trend Data Analysis 356
  15.1.1 Primary Trait Scoring of the Writing Tasks and Measures of Scorer Reliability 357
  15.1.2 Item Analysis of the Writing Trend Items 358
  15.1.3 Calibration of Writing Trend Prompts Using the Generalized Partial Credit Model 358
  15.1.4 Linking the Trend Data and Generating Plausible Values 362
  15.1.5 Other Analysis of Trends in Writing Performance 364
15.2 Main Data Analysis 366
  15.2.1 Item Analysis of the Main Assessment Prompts 366
  15.2.2 Calibration of Main Assessment Prompts and Generation of Plausible Values 366
  15.2.3 Linking the Main Assessment Data 370
15.3 Writing Portfolio Data Analysis 371

PART III
Statistical Summary of 1992 NAEP Data

Chapter 16 Statistical Summary of 1992 NAEP Data and Estimates of the Proficiencies of American Students 375
Bruce A. Kaplan, Laura S. Jerry, and Phillip T. Leung

16.1 Measurement Instruments 376
16.2 Sample Characteristics 376
16.3 Population Estimates 377

Appendix A Consultants for the Development of the 1992 NAEP Objectives and Items 529
Appendix B Reporting Subgroups and Derived Variables Used for Reporting and Conditioning 539
Appendix C Distribution of Weight Components 563
Appendix D Design Effects Statistics 583
Appendix E Subscale Item Counts and IRT Parameters 611
Appendix F Conditioning Variables and Related Data 667
Appendix G Mathematics Achievement Level-setting 839
Appendix H Reading Achievement Level-setting 863
Appendix I Mathematics Scale Anchoring 891
Appendix J Reading Scale Anchoring 909
Appendix K Open-ended Item Score Statistics 927
Appendix L Differential Item Functioning (DIF) Results 951
Appendix M Re-analysis of the 1990 Mathematics Assessment Data 957
Appendix N Estimation Error Variance by Gender and Race/ethnicity 967
Appendix O Correspondence Related to the Analysis of the 1992 Mathematics Short-term Trend Data 985
Appendix P Grade 8 Writing Trend: Investigation of the Changes in Mean Proficiency Between 1988 and 1990, and Between 1990 and 1992 999
References Cited in Text 1021

THE NAEP 1992 TECHNICAL REPORT

LIST OF TABLES

Table 1-1 NAEP 1992 student samples 12
1-2 NAEP subject areas, grades, and ages assessed: 1969-1992 16
1-3 Main assessment mathematics booklet design 25
1-4 Main assessment reading booklet design, age 13/grade 8 and age 17/grade 12 26
1-5 Main assessment reading booklet design, age 9/grade 4 26

Table 2-1 Percentage distribution of items by grade and reading purpose 43
2-2 Percentage distribution of items by reading stance 43
2-3 Distribution of questions for the 1992 reading assessment 45
2-4 Target and actual percentage distribution of questions by grade and reading purpose 45
2-5 Target and actual percentage distribution of questions by grade and reading ability 46
2-6 NAEP mathematics framework: Percentage distribution of items by grade and ability 54
2-7 NAEP mathematics framework: Percentage distribution of items by grade and content area 54
2-8 Total number of cognitive questions in the 1992 mathematics assessment 56
2-9 Distribution of mathematics questions across content and ability areas 56

Table 3-1 Assessment type by age class and assessment period 68
3-2 Geographic regions used by stratification 69
3-3 Sampling subuniverses and the number of noncertainty strata in each 70
3-4 Grade definition of school eligibility for frame inclusion 72
3-5 School sample sizes, refusals, and substitutes for the main samples 75
3-6 School sample sizes, refusals, and substitutes for the long-term trend samples 78
3-7 Number of students per school for each session type 81
3-8 Student exclusion rates by age class and school type, weighted 82
3-9 Student participation rates by age class and school type, unweighted 83
3-10 Overall participation rates by age class, unweighted 84
3-11 Weighted participation rates by age class and session type, main samples 85
3-12 Weighted participation rates by age class and session type, long-term trend samples 85

Table 4-1 Main sample booklet contents and number of booklets administered, age 9/grade 4 reading 89
4-2 Cognitive and noncognitive block information, age 9/grade 4 reading 89
4-3 Main sample booklet contents and number of booklets administered, age 13/grade 8 reading 90
4-4 Cognitive and noncognitive block information, age 13/grade 8 reading 90
4-5 Main sample booklet contents and number of booklets administered, age 17/grade 12 reading 91
4-6 Cognitive and noncognitive block information, age 17/grade 12 reading 91
4-7 Main sample booklet contents and number of booklets administered, age 9/grade 4 mathematics 93
4-8 Main sample booklet contents and number of booklets administered, age 13/grade 8 mathematics 94
4-9 Main sample booklet contents and number of booklets administered, age 17/grade 12 mathematics 95
4-10 Cognitive and noncognitive block information, age 9/grade 4 mathematics 96

4-11 Cognitive and noncognitive block information, age 13/grade 8 mathematics 97
4-12 Cognitive and noncognitive block information, age 17/grade 12 mathematics 98
4-13 Main sample booklet contents and number of booklets administered, age 9/grade 4 writing 101
4-14 Cognitive and noncognitive block information, age 9/grade 4 writing 101
4-15 Main sample booklet contents and number of booklets administered, age 13/grade 8 writing 102
4-16 Cognitive and noncognitive block information, age 13/grade 8 writing 102
4-17 Main sample booklet contents and number of booklets administered, age 17/grade 12 writing 103
4-18 Cognitive and noncognitive block information, age 17/grade 12 writing 103
4-19 Long-term trend sample booklet contents and number of booklets administered, age class 9 104
4-20 Long-term trend sample booklet contents and number of booklets administered, age class 13 104
4-21 Long-term trend sample booklet contents and number of booklets administered, age class 17 105
4-22 Long-term trend sample block information, age class 9 105
4-23 Long-term trend sample block information, age class 13 106
4-24 Long-term trend sample block information, age class 17 107

Table 5-1 Summary of field operations and data collection schedule 114
5-2 Session summary, long-term trend assessments 118
5-3 Session summary, main assessments 125
5-4 Completed questionnaires returned by NAEP supervisors to NCS 128

Table 7-1 Numbers of constructed-response questions 143
7-2 Reader percent agreement 154

Table 8-1 Summary of quality control error analysis for NAEP 1992 data entry 163

Table 10-1 Distribution of populations of eligible students based on full weighted sample of eligible schools, main samples 202
10-2 Distribution of populations of eligible students based on weighted sample of participating schools with school nonresponse adjustments, main samples 203
10-3 Distribution of populations of eligible students based on assessed and absent students from participating schools, main samples 204
10-4 Distribution of populations of eligible students based on assessed students from participating schools with student nonresponse adjustments applied, main samples 205
10-5 Distribution of populations of eligible students in participating schools based on invited students from participating schools using base weights unadjusted for school nonresponse, main writing samples 207
10-6 Distribution of populations of eligible students based on invited students in participating schools using weights adjusted for school nonresponse but not student nonresponse, main writing samples 208
10-7 Distributions of populations of eligible students based on assessed students from participating schools with students nonresponse adjustments applied, main writing samples 209

10-8 Weighted distribution of absent students by nature of absenteeism for all age/grades, writing assessment 210
10-9 Distribution of populations of eligible students based on trimmed weights of assessed students in participating schools, main writing samples 212
10-10 Major subgroups for poststratification in 1992 213
10-11 Distributions of design effects by demographic subgroups for proportion-correct statistics averaged across main samples, grade 4 224
10-12 Distributions of design effects by demographic subgroups for proportion-correct statistics averaged across main samples, grade 8 225
10-13 Distributions of design effects by demographic subgroups for proportion-correct statistics averaged across main samples, grade 12 226
10-14 Distributions of design effects by demographic subgroups for proportion-correct statistics averaged across long-term trend reading and writing samples, age 9 227
10-15 Distributions of design effects by demographic subgroups for proportion-correct statistics averaged across long-term trend reading and writing samples, age 13 228
10-16 Distributions of design effects by demographic subgroups for proportion-correct statistics averaged across long-term trend reading and writing samples, age 17 229
10-17 Distributions of design effects by demographic subgroups for proportion-correct statistics averaged across long-term trend mathematics and science samples, age 9 230
10-18 Distributions of design effects by demographic subgroups for proportion-correct statistics averaged across long-term trend mathematics and science samples, age 13 231
10-19 Distributions of design effects by demographic subgroups for proportion-correct statistics averaged across long-term trend mathematics and science samples, age 17 232
10-20 Distributions of design effects by demographic subgroups for proportion-correct statistics averaged across state samples, mathematics grade 4 233
10-21 Distributions of design effects by demographic subgroups for proportion-correct statistics averaged across state samples, mathematics grade 8 234
10-22 Distributions of design effects by demographic subgroups for proportion-correct statistics averaged across state samples, reading grade 4 235

Table 11-1 Estimation error variance and related coefficients for the 1992 grade 4 reading composite (based on five plausible values) 255

Table 12-1 NAEP 1992 reading student samples 258
12-2 NAEP reading samples contributing to 1992 long-term trend results 260
12-3 Numbers of scaled reading long-term trend items common across ages 261
12-4 Numbers of scaled reading long-term trend items common across assessments 261
12-5 Descriptive statistics for item blocks, reading long-term trend samples 263
12-6 Items deleted from the reading long-term trend analysis 264
12-7 Items calibrated separately by assessment year in the reading long-term trend analysis 266
12-8 Proportion of proficiency variance accounted for by the conditioning model for the reading long-term trend assessment 269
12-9 Means and standard deviations on the reading long-term trend proficiency scale 270
12-10 Estimation error variance and related coefficients for the reading long-term trend assessment 271

12-11 Item type composition of blocks used in reading main assessment 274
12-12 Numbers of scaled reading items common across grade levels 275
12-13 Descriptive statistics for item blocks by position within booklet and over all occurrences, reading main sample, grade 4 276
12-14 Descriptive statistics for item blocks by position within booklet and over all occurrences, reading main sample, grade 8 277
12-15 Descriptive statistics for item blocks by position within booklet and over all occurrences, reading main sample, grade 12 278
12-15 DIF category by grade 281
12-17 Recoding of polytomous items for scaling 283
12-18 Items scaled separately by age/grade 286
12-19 Proportion of proficiency variance accounted for by the conditioning model for the reading main assessment 293
12-20 Correlations and variances from conditioning 293
12-21 Correlations of reading scales 294
12-22 Coefficients of linear transformations of the purposes-of-reading scales from the calibrating scale units to the units of reporting proficiency 295
12-23 Weights used for each purpose-of-reading scale to form the reading composite 296
12-24 Means and standard deviations on the reading composite scale 296
12-25 Estimation error variance and related coefficients for the reading main assessment 297

Table 13-1 NAEP 1992 mathematics student samples 300
13-2 NAEP mathematics samples contributing to 1992 long-term trend results 302
13-3 Numbers of scaled mathematics long-term trend items common across ages 303
13-4 Numbers of scaled mathematics long-term trend items common across assessments 303
13-5 Correspondence between samples, booklets, and blocks for the mathematics long-term trend 304
13-6 Descriptive statistics for item blocks, mathematics long-term trend samples 305
13-7 Items deleted from the age 9 mathematics long-term trend analysis 307
13-8 Items deleted from the age 13 mathematics long-term trend analysis 307
13-9 Items deleted from the age 17 mathematics long-term trend analysis 308
13-10 Means and standard deviations on the mathematics long-term trend proficiency scale 310
13-11 Estimation error variance and related coefficients for the mathematics long-term trend assessment 310
13-12 Main focused-BIB: Numbers of scaled mathematics items common across grade levels, by content area scale 312
13-13 Main focused-BIB: Numbers of items in various formats, by grade 313
13-14 Difference in mean proportion correct for blocks in the grade 4 calculator bridge and the 1992 main assessment 314
13-15 Blocks used in the scaling of the 1992 mathematics assessment 314
13-16 Number of cognitive items in the 1992 special mathematics assessment 316
13-17 Percentage of students with all items incorrect or omitted for complex problem-solving blocks, by grade and race/ethnicity 317
13-18 Number of estimation and complex problem-solving items common across grades 317
13-19 Descriptive statistics for item blocks by position within booklet and over all occurrences, grade 4 319
13-20 Descriptive statistics for item blocks by position within booklet and over all occurrences, grade 8 320
13-21 Descriptive statistics for item blocks by position within booklet and over all occurrences, grade 12 321
13-22 Descriptive statistics for the estimation block by grade 322

13-23 Mathematics DIF category by grade 324
13-24 Items from the 1992 mathematics main analysis receiving special treatment 326
13-25 IRT scaling results for the estimation scale: Items receiving special treatment 325
13-26 Proportion of variance accounted for by the conditioning model for the mathematics main assessment 331
13-27 Coefficients of the linear transformations that transform the five content area scales from the 1992 calibration metric to the 1990 reporting proficiency metric 333
13-28 Coefficients of the linear transformations that transform the estimation scale from the 1992 calibration metric to the 1990 reporting proficiency metric 333
13-29 Defining weights for the mathematics composite by age/grade 334
13-30 Means and standard deviations on the mathematics composite scale 334
13-31 Average percent correct for the 1992 assessment versus a comparable subset of the 1990 assessment 335
13-32 Results of partitioning analysis for change between 1992 and 1990 in unweighted percent of items correct in a block, averaged across blocks 336
13-33 Results of partitioning analysis for change between 1992 and 1990 in weighted percent of items correct in a block, averaged across blocks 336
13-34 Estimation error variance and related coefficients for the mathematics main assessment 338
13-35 Conditional and marginal proficiency correlations, grade 4 338
13-36 Conditional and marginal proficiency correlations, grade 8 339
13-37 Conditional and marginal proficiency correlations, grade 12 339
13-38 Grade 4: Comparing a four-scale composite with a five-scale composite at the mean and selected percentiles 342

Table 14-1 NAEP 1992 science student samples 343
14-2 NAEP science samples contributing to 1992 long-term trend results 345
14-3 Numbers of scaled science long-term trend items common across ages 346
14-4 Numbers of scaled science long-term trend items common across assessments 346
14-5 Descriptive statistics for item blocks, science long-term trend samples 358
14-6 Correspondence between samples, booklets, and blocks for the science long-term trend 359
14-7 Items deleted from the science paced-tape long-term trend analysis 359
14-8 Proportion of proficiency variance accounted for by the conditioning model for the science long-term trend assessment 350
14-9 Means and standard deviations on the science long-term trend proficiency scale 352
14-10 Estimation error variance and related coefficients for the science long-term trend assessment 353

Table 15-1 NAEP 1992 writing student samples 356
15-2 Assignment of 1984-1992 writing long-term trend items in 1984, 1988, 1990, and 1992 357
15-3 Percentages of exact score agreement and interrater reliability for the primary trait scoring of the writing long-term trend assessment items 359
15-4 Descriptive statistics for writing prompts, writing long-term trend samples 360
15-5 Sample sizes for scaling items used in trends in writing performance 361
15-6 Linear transformation constants for the writing long-term trend between-grade linking 363
15-7 Means and standard deviations on the writing long-term trend proficiency scale 364
15-8 Estimation error variance and related coefficients for the writing long-term trend assessment 364
15-9 Sample sizes for mechanics scoring 365
15-10 Sample sizes for holistic scoring 365
15-11 Item number, purpose of writing, and block number for each writing prompt 367

15-12 Percentages of exact score agreement and interrater reliability for the primary trait scoring of the writing main assessment items 368
15-13 Descriptive statistics for writing prompts, writing main samples 369
15-14 Linear transformation constants for the writing main between-grade linking 370
15-15 Estimation error variance and related coefficients for the writing main assessment 371

Table 16-1 Measurement instruments used in 1992 NAEP 380
16-2 Number of items administered, by age class 381
16-3 Number of cognitive items administered by sample and age class 382
16-4 Characteristics of schools in main samples 383
16-5 Characteristics of schools in long-term trend samples 385
16-6 Numbers of responses to teacher questionnaires and students matched with teacher data 387
16-7 Numbers of assessment sessions by sample, type of session, and age class 388
16-8 Number of students assessed and excluded by sample and age class 389
16-9 Number of students in reading main sample by type of eligibility and subgroup classification, age 9/grade 4 390
16-10 Number of students in mathematics main sample by type of eligibility and subgroup classification, age 9/grade 4 391
16-11 Number of students in mathematics estimation sample by type of eligibility and subgroup classification, age 9/grade 4 392
16-12 Number of students in mathematics calculator sample by type of eligibility and subgroup classification, age 9/grade 4 393
16-13 Number of students in writing main sample by type of eligibility and subgroup classification, age 9/grade 4 394
16-14 Number of students in reading main sample by type of eligibility and subgroup classification, age 13/grade 8 395
16-15 Number of students in mathematics main sample by type of eligibility and subgroup classification, age 13/grade 8 396
16-16 Number of students in mathematics estimation sample by type of eligibility and subgroup classification, age 13/grade 8 397
16-17 Number of students in writing main sample by type of eligibility and subgroup classification, age 13/grade 8 398
16-18 Number of students in reading main sample by type of eligibility and subgroup classification, age 17/grade 12 399
16-19 Number of students in mathematics main sample by type of eligibility and subgroup classification, age 17/grade 12 400
16-20 Number of students in mathematics estimation sample by type of eligibility and subgroup classification, age 17/grade 12 401
16-21 Number of students in writing main sample by type of eligibility and subgroup classification, age 17/grade 12 402
16-22 Number of students in reading and writing long-term trend sample by type of eligibility and subgroup classification, age 9/grade 4 403
16-23 Number of students in reading and writing long-term trend sample by type of eligibility and subgroup classification, age 13/grade 8 404
16-24 Number of students in reading and writing long-term trend sample by type of eligibility and subgroup classification, age 17/grade 11 405
16-25 Number of students in mathematics and science long-term trend sample by type of eligibility and subgroup classification, age 9 406
16-26 Number of students in mathematics and science long-term trend sample by type of eligibility and subgroup classification, age 13 407

16-27 Number of students in mathematics and science long-term trend sample by type of eligibility and subgroup classification, age 17 408
16-28 Number of excluded students in main sample by type of eligibility and subgroup classification, age 9/grade 4 409
16-29 Number of excluded students in main sample by type of eligibility and subgroup classification, age 13/grade 8 410
16-30 Number of excluded students in main sample by type of eligibility and subgroup classification, age 17/grade 12 411
16-31 Number of excluded students in long-term trend samples by type of eligibility and subgroup classification, age 9/grade 4 412
16-32 Number of excluded students in long-term trend samples by type of eligibility and subgroup classification, age 13/grade 8 413
16-33 Number of excluded students in long-term trend samples by type of eligibility and subgroup classification, age 17/grade 12 414
16-34 Number of students by sample and age class 415
16-35 Number of students assessed in reading main sample, age 9/grade 4 416
16-36 Number of students assessed in mathematics main sample, age 9/grade 4 417
16-37 Number of students assessed in mathematics estimation sample, age 9/grade 4 418
16-38 Number of students assessed in mathematics calculator bridge sample, age 9/grade 4 419
16-39 Number of students assessed in writing main sample, age 9/grade 4 420
16-40 Number of students assessed in reading main sample, age 13/grade 8 421
16-41 Number of students assessed in mathematics main sample, age 13/grade 8 422
16-42 Number of students assessed in mathematics estimation sample, age 13/grade 8 423
16-43 Number of students assessed in writing main sample, age 13/grade 8 424
16-44 Number of students assessed in reading main sample, age 17/grade 12 425
16-45 Number of students assessed in mathematics main sample, age 17/grade 12 426
16-46 Number of students assessed in mathematics estimation sample, age 17/grade 12 427
16-47 Number of students assessed in writing main sample, age 17/grade 12 428
16-48 Number of students assessed in reading and writing long-term trend sample, age 9/grade 4 429
16-49 Number of students assessed in reading and writing long-term trend sample, age 13/grade 8 430
16-50 Number of students assessed in reading and writing long-term trend sample, age 17/grade 11 431
16-51 Number of students assessed in mathematics and science long-term trend samples 432
16-52 Weighted percentage of students in reading main sample by type of eligibility and subgroup classification, age 9/grade 4 433
16-53 Weighted percentage of students in mathematics main sample by type of eligibility and subgroup classification, age 9/grade 4 434
16-54 Weighted percentage of students in mathematics estimation sample by type of eligibility and subgroup classification, age 9/grade 4 435
16-55 Weighted percentage of students in mathematics calculator sample by type of eligibility and subgroup classification, age 9/grade 4 436
16-56 Weighted percentage of students in writing main sample by type of eligibility and subgroup classification, age 9/grade 4 437
16-57 Weighted percentage of students in reading main sample by type of eligibility and subgroup classification, age 13/grade 8 438
16-58 Weighted percentage of students in mathematics main sample by type of eligibility and subgroup classification, age 13/grade 8 439
16-59 Weighted percentage of students in mathematics estimation sample by type of eligibility and subgroup classification, age 13/grade 8 440

16-60 Weighted percentage of students in writing main sample by type of eligibility and subgroup classification, age 13/grade 8 441
16-61 Weighted percentage of students in reading main sample by type of eligibility and subgroup classification, age 17/grade 12 442
16-62 Weighted percentage of students in mathematics main sample by type of eligibility and subgroup classification, age 17/grade 12 443
16-63 Weighted percentage of students in mathematics estimation sample by type of eligibility and subgroup classification, age 17/grade 12 444
16-64 Weighted percentage of students in writing main sample by type of eligibility and subgroup classification, age 17/grade 12 445
16-65 Weighted percentage of students in reading and writing long-term trend sample by type of eligibility and subgroup classification, age 9/grade 4 446
16-66 Weighted percentage of students in reading and writing long-term trend sample by type of eligibility and subgroup classification, age 13/grade 8 447
16-67 Weighted percentage of students in reading and writing long-term trend sample by type of eligibility and subgroup classification, age 17/grade 11 448
16-68 Weighted percentage of students in mathematics and science long-term trend sample by type of eligibility and subgroup classification, age 9 449
16-69 Weighted percentage of students in mathematics and science long-term trend sample by type of eligibility and subgroup classification, age 13 450
16-70 Weighted percentage of students in mathematics and science long-term trend sample by type of eligibility and subgroup classification, age 17 451
16-71 Weighted percentage of excluded students in main samples by type of eligibility and subgroup classification, age 9/grade 4 452
16-72 Weighted percentage of excluded students in main samples by type of eligibility and subgroup classification, age 13/grade 8 453
16-73 Weighted percentage of excluded students in main samples by type of eligibility and subgroup classification, age 17/grade 12 454
16-74 Weighted percentage of excluded students in long-term trend samples by type of eligibility and subgroup classification, age 9/grade 4 455
16-75 Weighted percentage of excluded students in long-term trend samples by type of eligibility and subgroup classification, age 13/grade 8 456
16-76 Weighted percentage of excluded students in long-term trend samples by type of eligibility and subgroup classification, age 17/grade 11 457
16-77 Number of students in NAEP main samples with proficiency scores by type of eligibility 458
16-78 Number of students in NAEP long-term trend samples with proficiency scores by type of eligibility 459
16-79 Weighted means, standard deviations, and percentiles of reading distributions for reading main samples, by subgroup, grade 4 460
16-80 Weighted means, standard deviations, and percentiles of reading distributions for reading main samples, by subgroup, grade 8 461
16-81 Weighted means, standard deviations, and percentiles of reading distributions for reading main samples, by subgroup, grade 12 462
15-82 Percentiles for reading long-term trend, grade 4 463
16-83 Percentiles for reading long-term trend, grade 8 465
16-84 Percentiles for reading long-term trend, grade 12 467
16-85 Weighted means, standard deviations, and percentiles of mathematics distributions for mathematics main samples, by subgroup, grade 4 469
16-86 Weighted means, standard deviations, and percentiles of mathematics distributions for mathematics main samples, by subgroup, grade 8 470

16-87 Weighted means, standard deviations, and percentiles of mathematics distributions for mathematics main samples, by subgroup, grade 12 471
16-88 Percentiles for mathematics long-term trend, grade 4 472
16-89 Percentiles for mathematics long-term trend, grade 8 474
16-90 Percentiles for mathematics long-term trend, grade 12 476
16-91 Percentiles for science long-term trend, grade 4 478
16-92 Percentiles for science long-term trend, grade 8 480
16-93 Percentiles for science long-term trend, grade 12 482
16-94 Weighted means, standard deviations, and percentiles of writing distributions for writing main samples, by subgroup, grade 4 484
16-95 Weighted means, standard deviations, and percentiles of writing distributions for writing main samples, by subgroup, grade 8 485
16-96 Weighted means, standard deviations, and percentiles of writing distributions for writing main samples, by subgroup, grade 12 486
16-97 Percentiles for writing long-term trend, grade 4 487
16-98 Percentiles for writing long-term trend, grade 8 488
16-99 Percentiles for writing long-term trend, grade 12 489
16-100 Weighted response percentages and proficiency means (with standard errors) for reading main samples, by gender, grade 4 490
16-101 Weighted response percentages and proficiency means (with standard errors) for reading main samples, by derived race/ethnicity, grade 4 491
16-102 Weighted response percentages and proficiency means (with standard errors) for reading main samples, by parents' education level, grade 4 492
16-103 Weighted response percentages and proficiency means (with standard errors) for reading main samples, by gender, grade 8 493
16-104 Weighted response percentages and proficiency means (with standard errors) for reading main samples, by derived race/ethnicity, grade 8 494
16-105 Weighted response percentages and proficiency means (with standard errors) for reading main samples, by parents' education level, grade 8 495
16-106 Weighted response percentages and proficiency means (with standard errors) for reading main samples, by gender, grade 12 496
16-107 Weighted response percentages and proficiency means (with standard errors) for reading main samples, by derived race/ethnicity, grade 12 497
16-108 Weighted response percentages and proficiency means (with standard errors) for reading main samples, by parents' education level, grade 12 498
16-109 Proficiencies for reading long-term trend samples, grade 4 499
16-110 Proficiencies for reading long-term trend samples, grade 8 500
16-111 Proficiencies for reading long-term trend samples, grade 12 501
16-112 Weighted response percentages and proficiency means (with standard errors) for mathematics main samples, by gender, grade 4 502
16-113 Weighted response percentages and proficiency means (with standard errors) for mathematics main samples, by derived race/ethnicity, grade 4 503
16-114 Weighted response percentages and proficiency means (with standard errors) for mathematics main samples, by parents' education level, grade 4 504
16-115 Weighted response percentages and proficiency means (with standard errors) for mathematics main samples, by gender, grade 8 505
16-116 Weighted response percentages and proficiency means (with standard errors) for mathematics main samples, by derived race/ethnicity, grade 8 506
16-117 Weighted response percentages and proficiency means (with standard errors) for mathematics main samples, by parents' education level, grade 8 507

16-118 Weighted response percentages and proficiency means (with standard errors) for mathematics main samples, by gender, grade 12 508
16-119 Weighted response percentages and proficiency means (with standard errors) for mathematics main samples, by derived race/ethnicity, grade 12 509
16-120 Weighted response percentages and proficiency means (with standard errors) for mathematics main samples, by parents' education level, grade 12 510
16-121 Proficiencies for mathematics long-term trend samples, grade 4 511
16-122 Proficiencies for mathematics long-term trend samples, grade 8 512
16-123 Proficiencies for mathematics long-term trend samples, grade 12 513
16-124 Proficiencies for mathematics long-term trend samples, age 9 514
16-125 Proficiencies for mathematics long-term trend samples, age 13 515
16-126 Proficiencies for mathematics long-term trend samples, age 17 516
16-127 Weighted response percentages and proficiency means (with standard errors) for writing main samples, by gender, grade 4 517
16-128 Weighted response percentages and proficiency means (with standard errors) for writing main samples, by derived race/ethnicity, grade 4 518
16-129 Weighted response percentages and proficiency means (with standard errors) for writing main samples, by parents' education level, grade 4 519
16-130 Weighted response percentages and proficiency means (with standard errors) for writing main samples, by gender, grade 8 520
16-131 Weighted response percentages and proficiency means (with standard errors) for writing main samples, by derived race/ethnicity, grade 8 521
16-132 Weighted response percentages and proficiency means (with standard errors) for writing main samples, by parents' education level, grade 8 522
16-133 Weighted response percentages and proficiency means (with standard errors) for writing main samples, by gender, grade 12 523
16-134 Weighted response percentages and proficiency means (with standard errors) for writing main samples, by derived race/ethnicity, grade 12 524
16-135 Weighted response percentages and proficiency means (with standard errors) for writing main samples, by parents' education level, grade 12 525
16-136 Proficiencies for writing long-term trend samples, grade 4 526
16-137 Proficiencies for writing long-term trend samples, grade 8 527
16-138 Proficiencies for writing long-term trend samples, grade 11 528

THE NAEP 1992 TECHNICAL REPORT

LIST OF FIGURES

Figure 2-1 Description of reading stances 41
2-2 Description of purposes for reading 42
2-3 Description of mathematical abilities 52
2-4 Description of mathematical content areas 53
2-5 NAEP primary trait scoring guidelines 62

Figure 5-1 Instructions for excluding students (long-term trend assessments) 121
5-2 Instructions for excluding students (main assessments) 126

Figure 6-1 Data flow overview 133

Figure 7-1 Outline for scoring of constructed responses for mathematics 145
7-2 Outline for scoring of constructed responses for reading 147
7-3 Outline for scoring writing responses 151

Figure 12-1 Example long-term trend item demonstrating DIF across assessment years 1990 and 1992 267
12-2 Example long-term trend item, age 17, fitting separate items response functions for each assessment year 268
12-3 Example cross-sectional polytomous item demonstrating poor model fit 284
12-4 Example cross-sectional polytomous item after collapsing categories 0 and 1 285
12-5 Example cross-sectional dichotomous item demonstrating DIF across age/grade samples 287
12-6 Example cross-sectional item fitting separate item response functions for each age/grade sample 288
12-7 Example cross-sectional polytomous item demonstrating DIF across age/grade samples 289
12-8 Example cross-sectional polytomous item fitting separate item response functions for each age/grade sample, age 9/grade 4 290
12-9 Example cross-sectional polytomous item fitting separate item response functions for each age/grade sample, age 13/grade 8 291

Figure 13-1 Example of poor item fit: One group 328
13-2 Example of poor item fit: Two subgroups 329
13-3 Examples of fitting a previously poor-fitting item to two separate subgroups 330

Figure 14-1 Plot comparing empirical and model-based estimates of the item response function for item N406303 in the 1992 long-term science trend samples 351

ACKNOWLEDGMENTS

The design, development, administration, analysis, and reporting of the 1992 National Assessment of Educational Progress (NAEP) was a collaborative effort among staff from the National Assessment Governing Board (NAGB), the National Center for Education Statistics (NCES), the Council of Chief State School Officers (CCSSO), Educational Testing Service (ETS), Westat, and National Computer Systems (NCS). This report documents the NAEP design, administration, and data analysis procedures and provides insight about the rationale behind the technical decisions made. The development of this report and of the national assessment program is the result of the considerable knowledge, experience, creativity, and dedication of many individuals. It is fitting that these individuals be acknowledged for their contributions to NAEP.

The 1992 NAEP was funded through the National Center for Education Statistics in the Office of Educational Research and Improvement of the U.S. Department of Education. Emerson Elliott, NCES commissioner, provided consistent support and guidance. The staff worked closely and collegially with ETS, Westat, and NCS staff and played a crucial role in all aspects of the program. We are particularly grateful for the contributions of Gary Phillips, Susan Ahmed, Peggy Carr, Pat Dabbs, Stephen Gorman, Andrew Kolstad, Larry Ogle, Eugene Owen, Mary Rollefson, Sharif Shikrani, Maureen Treacy, and Sheida White. We would also like to thank the anonymous reviewers arranged by NCES for their helpful comments. The members of the National Assessment Governing Board (NAGB) and the NAGB staff also provided advice and guidance.

ETS management and the Center for the Assessment of Educational Progress (CAEP) have been very supportive of NAEP's technical work. Special thanks go to the late Gregory Anrig and to Nancy Cole as well as to Henry Braun and Charles Davis of ETS research management, and Archie Lapointe, Ina Mullis, Jules Goodison, Paul Williams, and David Hobson of CAEP. Significant contributions to the project were also received from Kent Ashworth, Alexandra Beatty, Jay Campbell, Mary Foertsch, Claudia Gentile, Janet Johnson, Chancey Jones, Lee Jones, Jeffrey Haberstroh, Lynn Jenkins, Stephen Koffler, Andrew Latham, Steve Lazer, Rebekkah Melchor-Logan, John Olson, Christine O'Sullivan, Hilary Persky, and Clyde Reese.

The guidance of the NAEP Design and Analysis Committee on technical aspects of NAEP has been outstanding. The members are Sylvia Johnson (chair), Albert Beaton, Jeri Benson, Clifford Clogg, Jeremy Finn, Huynh Huynh, David Lohman, Bengt Muthén, Anthony Nitko, Ingram Olkin, Tej Pandey, and Juliet Shaffer. Former committee members who were also advisors for the 1992 assessment were John Carroll, William Cooley, and Bert Green.

The design and data analysis of the 1992 National Assessment were primarily the responsibility of the NAEP research and data analysis staff with significant contributions from NAEP management, Westat, and NCS staffs. In addition to managing day-to-day data analytic operations, these staff members have made many innovative statistical and psychometric contributions. Major contributions were made by Nancy Allen, James Carlson, Huahua Chang, John Donoghue, Angela Grima, Frank Jenkins, John Mazzeo, Eiji Muraki, Spencer Swinton, and Neal Thomas. Jo-lin Liang was exceptional in her role as technical assistant. Robert Mislevy and Ming-mei Wang provided valuable statistical and psychometric advice.

The division of Data Analysis and Technology Research, under the outstanding leadership of John Barone, was responsible for developing the operating systems and carrying out the data analyses. Alfred Rogers and David Freund deserve special recognition for their leadership in developing and maintaining the large and complex NAEP data management systems. Alfred Rogers also deserves special mention for his role in the development of production versions of key analysis and scaling systems. Special thanks also go to David Freund, Steven Isham, Bruce Kaplan, and Edward Kulick for their continuing roles as leaders and developers of innovative software solutions to NAEP data analysis challenges. Other members of this division who made substantial contributions of their talent, and important contributions to NAEP data analyses, were Drew Bowker, Lucie Chan, John Ferris, Yim Fai Fong, Laura Jerry, Phillip Leung, Laura Jenkins, Michael Narcowich, Jennifer Nelson, Inge Novatkoski, Robert Patrick, Craig Pizzuti, Ira Sample, and Minhwei Wang.

The staff of Westat, Inc. contributed their talents and efforts in all areas of the sample design and data collection. Particular recognition is due to Renee Slobasky and Nancy Caldwell for supervising the field operations, to Keith Rust for developing and supervising the sampling design and selection, and to Debra Vivari for supervising the data processing operations. Thanks are also due to Bridgett Bell, Annell Bond, Frances Cohen, Lesly Flemming, Dward Moore, Sandra Rieder, Marilyn Rowen, Lana Ryaboy, Valerija Smith, Mark Waksberg, Leslie Wallace, Dianne Walsh, and Marianne Whitlock. The exceptional data collection effort of the Westat field staff is also gratefully acknowledged.

Critical to the program was the contribution of National Computer Systems, Inc., which has been responsible for the printing, distribution, and processing of the assessment materials. The leadership roles of John O'Neill and Judith Moyer are especially acknowledged. Thanks go also to Linda Reynolds, Bradley Thayer, Dianne Smrdel, Lavonne Mohn, and Mathilde Kennel.

Judy Alfort, Donna Lembeck, and Mary Varone are acknowledged for their patience and diligence in typing and proofing the many revisions of this report. Kent Ashworth was responsible for coordinating the design of the report's cover, which was created by Joyce Hofstetter.

Special thanks go to Debra Kline, the editor of this report, for organizing, scheduling, editing, motivating, and ensuring the cohesiveness and correctness of the final report.

Finally, NAEP is grateful to the students and school staff members who participated in the assessment. Without their efforts, there would be no assessment.

Eugene G. Johnson
Coordinating Director
NAEP Research
July 15, 1994

INTRODUCTION

THE NAEP 1992 TECHNICAL REPORT

Introduction

Eugene G. Johnson and James E. Carlson

Educational Testing Service

The 1992 National Assessment of Educational Progress (NAEP) monitored the performance of students in American schools in the subject areas of reading, mathematics, science, and writing. The sample involved more than 145,000 public- and private-school students who were 9-, 13-, or 17 years old or in grades 4, 8, 11, or 12.

The purpose of this technical report is to provide details on the instrument development, sample design, data collection, and data analysis procedures of the 1992 assessment. Substantive results are not presented here but can be found in a series of NAEP reports on the status of and trends in student performance[1].

Additional samples of approximately 100,000 fourth-graders and 100,000 eighth-graders in public-schools in 44 states and territories were assessed in mathematics and another 100,000 fourth-grade students were assessed in reading as part of the 1992 Trial State Assessment. A representative sample of about 2,500 students for each grade and subject was selected in each jurisdiction. The state-level sampling plan allowed for cross-state comparisons and comparisons with the nation in fourth- and eighth-grade mathematics achievement and fourth-grade reading achievement. Technical details of the Trial State Assessment are not presented in this technical report but can be found in the *Technical Report of the NAEP 1992 Trial State Assessment Program in Mathematics* (Johnson, Mazzeo, & Kline, 1993), and the *Technical Report of the NAEP 1992 Trial State Assessment Program in Reading* (Johnson, Mazzeo, & Kline, 1994).

An Overview of NAEP in 1992

For the 1992 assessment, NAEP researchers continued to build on the original design technology outlined in *A New Design for a New Era* (Messick, Beaton, & Lord, 1983). In order to maintain its links to the past and still implement innovations in measurement technology, NAEP continued its two-tiered sampling approach. Trend samples use the same methodology and population definitions as in previous assessments. Main assessment samples use innovations associated with new NAEP technology and address current educational issues. Trend sample data are used to estimate changes in performance from previous assessments; main assessment sample data are used for analyses involving the current student population. In continuing to use this two-tiered approach, NAEP reaffirms its commitment to maintaining long-term trends while at the same time implementing the latest in measurement technology.

---

[1] *NAEP 1992 Reading Report Card for the Nation and the States* (Mullis, Campbell, & Farstrup, 1993); *NAEP 1992 Mathematics Report Card for the Nation and the States* (Mullis, Dossey, Owen, & Phillips, 1993); *NAEP 1992 Writing Report Card* (Applebee, Langer, Mullis, Latham, & Gentile, 1994); *NAEP 1992 Trends in Academic Progress: Achievement of U.S. Students in Science, 1969 to 1992; Mathematics, 1973 to 1992; Reading, 1971 to 1992; Writing, 1984 to 1992* (Mullis, Dossey, Campbell, Gentile, O'Sullivan, & Latham, 1994); and *Windows into the Classroom: NAEP's 1992 Writing Portfolio Study* (Gentile & Martin-Rehrmann, 1994).

In the 1992 assessment, many of the innovations that were implemented for the first time in 1988 were continued and enhanced. For example, the use of the focused balanced incomplete block (focused-BIB) booklet design that began in 1988 was continued in 1992 for the main assessment samples in reading, mathematics, and writing. In the focused-BIB design, an individual receives blocks of cognitive items in the same subject area. The focused-BIB design allows for improved estimation of composite scale scores and subscale scores within a particular subject area.

NAEP in 1992 continued to apply the plausible values approach to estimating means for demographic as well as curriculum-related subgroups. Proficiency estimates were based on "draws" from a posterior distribution that was based on an optimum weighting of two sets of information—the student's responses to cognitive items and his or her demographic and associated educational process variables. This Bayesian procedure was developed by Mislevy (see Chapter 11 or Mislevy, 1991). An improvement that was implemented first in 1988 and refined for the 1992 assessment is the multivariate procedure that uses information from all scales within a given subject area in the estimation of the proficiency distribution on any one scale in that subject area.

A major improvement in the 1992 assessment was the introduction of the generalized partial credit model for IRT scaling. This allowed the incorporation of constructed-response questions that are scored on a multipoint rating scale into the NAEP scale in a way that utilizes the information available in each response category.

One important innovation in reporting the 1992 assessment data that was initiated in 1990 was the use of Bonferroni multiple comparison procedures to form confidence intervals for the trend differences between the current assessment year and each previous assessment year. Methods such as the Bonferroni allow one to control for the type 1 error rate for a fixed number of comparisons. In addition to the Bonferroni procedures, tests for linear and quadratic trends were also applied to the national trend data in reading, mathematics, science, and writing. It is anticipated that future NAEP reports will continue to build on these developments incorporating more powerful multiple comparison methods that will more closely approach an optimal balance between power and family error rates in large tables.

Organization of the Technical Report

Part I of this report presents the details of the design of the 1992 National Assessment, summarized in Chapter 1. Chapters 2 through 8 describe the development of the objectives and the items used in the assessment, the sample selection procedures, the assessment booklets and questionnaires, the administration of the assessment in the field, the processing of the data from the assessment instruments into computer-readable form, the professional scoring of constructed-response items, and the methods used to create a complete NAEP database.

The 1992 NAEP data analysis procedures are described in Part II of the report. Chapter 9 provides a summary of the analysis steps. Subsequent chapters provide a general discussion of the weighting and variance estimation procedures used in NAEP, an overview of NAEP scaling methodology, and details of the trend and main assessment analyses performed for each subject area in the 1992 assessment.

Part III presents basic data from the 1992 assessment, including the properties of the measuring instruments, characteristics of the sample, and selected estimates of the proficiencies of students in each of the subject areas assessed.

PART I

The Design and Implementation of the 1992 NAEP

Chapter 1

OVERVIEW OF PART I: THE DESIGN AND IMPLEMENTATION OF THE 1992 NAEP[1]

Eugene G. Johnson

Educational Testing Service

The 1992 National Assessment collected information on the knowledge, skills, understanding, and attitudes of young Americans in reading, mathematics, science, and writing. The basis for this information was a complex sample survey involving more than 145,000 students, consisting of national samples of public- and private-school students who were aged 9, 13, and 17 or in grades 4, 8, 11, and 12. Additional data came from the Trial State Assessment Program, which in 1992 assessed mathematics at grades 4 and 8 and reading at grade 4 in representative samples of public-school students in 41 states, the District of Columbia, and two territories.

This chapter describes the design for the 1992 assessment and gives an overview of the steps involved in its implementation, from the planning stage through the creation of edited data files. The major components of the implementation process are presented here with references to the appropriate chapters in Part I for more details. The procedures used for the analysis of the data are summarized in the overview to Part II and discussed in detail in the remaining chapters in that part of the report. Excluded from this technical report are the details of the design and analysis of the 1992 Trial State Assessment, which instead appear in the *Technical Report of the NAEP 1992 Trial State Assessment Program in Mathematics* (Johnson, Mazzeo, & Kline, 1993) and the *Technical Report of the NAEP 1992 Trial State Assessment Program in Reading* (Mazzeo, Johnson, & Kline, 1994).

The organization of this chapter, and of Part I, is as follows:

- Section 1.1 provides an overview of the NAEP design for 1992 and describes the constituent samples. To provide background information, the section also gives the assessment schedule from the inception of NAEP in 1969 through the 1992 assessment.

- Section 1.2 summarizes the development of the objectives for each subject area in the assessment and describes the development and review of the items written to fit those objectives. Details of the objective and item development processes appear in Chapter 2.

---

[1] The author is indebted to the authors of Chapters 2 through 8 for portions of this chapter.

- Section 1.3 summarizes the four-stage stratified random sampling procedures used for the 1992 assessment with a more full description provided in Chapter 3.
- Section 1.4 discusses the assignment of the cognitive and background questions to assessment booklets and describes the focused-BIB spiral design. Chapter 4 provides a detailed description of the assessment booklets.
- Section 1.5 summarizes the field administration procedures, including the processes of training field administrators, attaining school cooperation, administering the assessment, and conducting quality control. Further details appear in Chapter 5.
- Section 1.6 describes the flow of data from the receipt of the assessment materials through data entry, validation, and resolution to the creation of edited data files. Chapter 6 provides a detailed description of the process.
- Section 1.7 discusses the professional scoring of students' responses to the open-ended items in the assessment. Details of the process are given in Chapter 7.
- Section 1.8 summarizes the creation of the database, the quality control of data entry, and lists the 1992 database products. Further details appear in Chapter 8.

1.1 THE 1992 NAEP DESIGN

A major charge to NAEP is to reliably measure trends over time in educational achievement. To do this well, confounding effects due to changes from one assessment to the next in assessment instrumentation or in assessment procedures must be minimized. This implies a stability in the measurement process over time. At the same time, the assessment must remain current by allowing the introduction of new curriculum concepts and changes in educational priorities and by permitting the use of new measurement technology. The objectives for an assessment are determined through a consensus process in which committees of subject matter experts, scholars, and citizens representing many diverse constituencies and points of view are assembled to determine the educational goals that students should achieve. Satisfying these objectives often requires changes in assessment instrumentation and methodology.

A solution to the dilemma of measuring trends while maintaining currency is to institute a multicomponent assessment system where each component is itself an assessment designed to accomplish a specific goal. There are three components in the 1992 design: (1) assessments for long-term trend; (2) main assessments; and (3) the Trial State Assessment. These are discussed in detail below.

A number of improvements were made in the design of NAEP in the 1984 and succeeding assessments. Until the 1984 assessment, NAEP was administered using matrix sampling and tape recorders; that is, by administering booklets of exercises using an aurally presented stimulus that paced the students through the individual assessment exercises. In the 1984 assessment, BIB spiraling, which does not include aural pacing, was introduced in place of taped matrix sampling. The NAEP design now includes sampling grade populations as well as the age populations that NAEP originally collected. To assure that the age/grade samples measure four years of growth, the definitions of student age and the time of year in which the assessment takes place have been made uniform. To shorten the timetable for reporting results, the period for national data collection was decreased in 1992 from the five-month period used in 1990 to a three-month period in the winter (corresponding to the period used for the winter half-sample of the 1990 National Assessment). To enhance the coverage of the subject areas assessed, the number of cognitive items administered was increased for the 1992 assessment.

The focus of assessments has also changed over time in response to changing priorities. For example, in contrast with previous assessments, the 1992 assessment of reading required many more open-ended responses and contained longer, more naturally occurring passages, including literary and informational texts as well as documents.

NAEP's design for 1992 required collecting 23 different types of samples in order to conduct the assessments and to monitor the modifications introduced into NAEP (such as the need to purchase new calculators at grade 4 for the Trial State Assessment). The various samples collected for the 1992 assessment are summarized in Table 1-1. Each row of Table 1-1 corresponds to a particular sample and each column of the table indicates the following major features of that sample:

1) *Sample* is the sample identifier. The first part of the sample code is a number representing the age class included in the sample; the second part, in brackets, denotes the specific sample type. For example, 9[Math-MainP] is a main assessment mathematics sample for age class 9. A full description of the purposes for the various sample types will be given below.

2) *Booklets* gives the identifier numbers for the booklets used for the assessment of the particular sample.

3) *Mode* indicates the mode of assessment, which may be print, tape, or oral. NAEP originally assessed students using a tape recorder in addition to booklets, thus pacing the students through exercises at a fixed rate. In 1992 NAEP used a paced audiotape for the mathematics assessment of estimation and for its long-term trend assessments. However, most other assessments in 1992 used printed instructions with the student expected to read the exercises. At age 9/grade 4, a national sample of students participated in an oral reading assessment administered to individual students in a one-to-one interview mode.

4) The *cohort assessed* denotes the age, grade, or age/grade of the population being sampled. For example, *age 9/grade 4* represents students who are *either* 9 years old *or* in the fourth grade; an *age 17* cohort consists of students (in any grade) who are 17 years old; a *grade 8* cohort consists of students (of any age) who are in the eighth grade. The traditional NAEP samples used in trend estimation were defined by age only. Populations for the 1992 main assessment were defined by being either of a particular age *or* of the modal grade for students of that age. The definitions of age, and thus the corresponding grade, have changed in ways that are described below.

Table 1-1
NAEP 1992 Student Samples

Sample	Booklets	Mode	Cohort Assessed	Time of Testing	Age Defn.	Modal Grade	Number Assessed
9[Math-MainP]	1 - 26	Print	Age 9/grade 4	1/6/92 - 3/31/92	CY	4	9,414
13[Math-MainP]	1 - 26	Print	Age 13/grade 8	1/6/92 - 3/31/92	CY	8	10,291
17[Math-MainP]	1 - 26	Print	Age 17/grade 12	1/6/92 - 3/31/92	CY	12	9,499
9[Math-MainT]	27	Tape	Age 9/grade 4	1/6/92 - 3/31/92	CY	4	2,054
13[Math-MainT]	27	Tape	Age 13/grade 8	1/6/92 - 3/31/92	CY	8	2,416
17[Math-MainT]	27	Tape	Age 17/grade 12	1/6/92 - 3/31/92	CY	12	2,074
9[Math-Calc]	28	Print	Age 9/grade 4	1/6/92 - 3/31/92	CY	4	2,236
9[Rdg-MainP]	30 - 45	Print	Age 9/grade 4	1/6/92 - 3/31/92	CY	4	8,416
13[Rdg-MainP]	30 - 49	Print	Age 13/grade 8	1/6/92 - 3/31/92	CY	8	14,942
17[Rdg-MainP]	30 - 50	Print	Age 17/grade 12	1/6/92 - 3/31/92	CY	12	15,315
9[Rdg-MainO]*	N/A	Oral	Grade 4	1/6/92 - 3/31/92	CY	4	1,583
9[Wrt-MainP]	60 - 77	Print	Age 9/grade 4	1/6/92 - 3/31/92	CY	4	9,552
13[Wrt-MainP]	60 - 79	Print	Age 13/grade 8	1/6/92 - 3/31/92	CY	8	14,942
17[Wrt-MainP]	60 - 80	Print	Age 17/grade 12	1/6/92 - 3/31/92	CY	12	15,669
9[Math-State]	1 - 26	Print	Grade 4	2/3/92 - 3/6/92	—	4	**
13[Math-State]	1 - 26	Print	Grade 8	2/3/92 - 3/6/92	—	8	**
9[Rdg-State]	30 - 45	Print	Grade 4	2/3/92 - 3/6/92	—	4	**
9[RW-LTTrend]	51 - 56	Print	Age 9/grade 4	1/6/92 - 3/31/92	CY	4	7,062
13[RW-LTTrend]	51 - 56	Print	Age 13/grade 8	10/7/91 - 12/13/91	CY	8	5,514
17[RW-LTTrend]	51 - 56	Print	Age 17/grade 11	3/16/92 - 5/15/92	Not CY	11	5,569
9[MS-LTTrend]	91 - 93	Tape	Age 9	1/6/92 - 3/31/92	CY	4	7,335
13[MS-LTTrend]	91 - 93	Tape	Age 13	10/7/91 - 12/13/91	CY	8	5,909
17[MS-LTTrend]	84 - 85	Tape	Age 17	3/16/92 - 5/15/92	Not CY	11	4,359

LEGEND:

Math	Mathematics
Rdg	Reading
Wrt	Writing
RW	Reading and writing
MS	Mathematics and science
MainP	Main assessment, print administration
LTTrend	Long-term trend assessment
Math-MainT	Mathematics estimation and problem solving with paced tape administration
Math-Calc	Bridge for new calculators at age 9/grade 4 in 1992
CY	Calendar year: birthdates in 1982, 1978, and 1974 respectively for ages 9, 13, and 17
Not CY	Age 17 only: birthdates between Oct. 1, 1974 and Sept. 30, 1975

* The 9[Rdg-MainO] students participating in the oral reading assessment are a subsample of those in the 9[Rdg-MainP] print assessment

** Consists of distinct samples in 44 states and jurisdictions

5) *Time of testing* indicates the time of year in which the assessment is performed. NAEP traditionally assessed 9-year-olds in the winter, 13-year-olds in the fall, and 17-year-olds in the spring; in the 1992 main assessment all age classes were assessed in the winter (between 1/6/92 and 3/31/92).

6) *Age definition* is denoted as calendar year (CY) or not calendar year (Not CY). NAEP originally defined age by birth within a calendar year at ages 9 and 13 but defined age 17 as being born between October 1 of one year and September 30 of the next. As in 1990, the 1992 main assessments placed all ages on the same calendar-year basis.

7) The *modal grade* is the grade attended by most of the students of the sampled age. For example, if an age 17 sample is listed as having a modal grade of 11, then most of the 17-year-old students, as defined, are in the eleventh grade. The definition of age affects the modal grade of the sample. The ages in samples in the 1992 main assessment were defined so that the modal grades were 4, 8, and 12.

8) The *number assessed* is the number of students in the sample who were actually administered the assessment.

1.1.1 The 1992 NAEP Samples

The NAEP samples in 1992 consisted of three types—the main NAEP samples, the long-term trend samples, and the Trial State Assessment samples.

The Main NAEP Samples. The main NAEP samples are labeled in Table 1-1 as [Math-MainP], [Math-MainT], [Math-Calc], [Rdg-MainP], [Rdg-MainO] and [Wrt-MainP]. The samples use the new assessment technology, focused-BIB spiraling (defined in section 1.4) or special innovative procedures, and are intended to form the basis for future assessments. Each age class sample is assessed in the winter period (1/6/92 to 3/31/92). In these samples, age is defined by calendar year, and both the age populations and modal grade populations are sampled. The various main NAEP samples, and their purposes, are as follows:

> [Math-MainP] are age/grade mathematics assessment samples used for measuring mathematics achievement in 1992 as well as for assessing short-term trends in mathematics performance from 1990. The fourth-and eighth-grade samples also provided the comparison groups for the 1992 Trial State Assessment of mathematics in grades 4 and 8. These samples used print administration.
>
> [Math-MainT] are samples used for assessing mathematical topics, such as estimation and complex problem-solving situations, that cannot be adequately measured using printed administration. These samples were administered using a tape recorder.
>
> 9[Math-Calc] is a sample administered at age 9/grade 4 that was designed to measure the effect of changing the type of calculator used. Because of the need

to purchase thousands of new calculators for the Trial State Assessment, a different calculator was used at age 9/grade 4 in 1992 than was used in 1990. The age 9/grade 4 students selected for the [Math-Calc] sample used the same calculator as was used in 1990. The effects of changing the type of calculator between 1990 and 1992 was bridged by comparing the responses of these students with the responses of the students in the 9[Math-MainP] sample. (Because the two older cohorts used the same calculator as was used in 1990, no bridge was necessary.)

> [Rdg-MainP] are age/grade reading assessment samples used for measuring reading achievement in 1992. The fourth-grade samples also provided the comparison groups for the 1992 Trial State Assessment of reading in grade 4.
>
> 9[Rdg-MainO] is a subsample of the students participating in the 9[Rdg-MainP] assessment. These students read aloud one of the passages in their print administration booklet and participated in a brief interview. They were tape recorded and their oral readings were analyzed for miscues.
>
> [Wrt-MainP] are age/grade writing assessment samples used for measuring writing achievement in 1992.

The Long-term Trend Samples. The long-term trend samples are labeled as [RW-LTTrend] and [MS-LTTrend] in Table 1-1. Each sample was defined in the same way as equivalent samples in previous assessments and used the same assessment technology as was used in those assessments. Therefore, the long-term trend samples are directly comparable to data from previous assessments and so can be used for continuing the NAEP long-term trend lines. Because these samples were designed to link the 1992 data with data from previous assessments, they are also referred to as bridge samples. The long-term trend samples and their purposes are as follows:

> [RW-LTTrend] are age/grade samples used for estimating long-term trends in reading and writing. These samples used assessment booklets identical to those initially used in 1984 and subsequently used in 1988 and 1990 (many of the items were also used in pre-1984 assessments). As in 1984, 1988,and 1990, print administration was used. These samples used the age definitions and time of testing originally used by NAEP in the 1970s and the early 1980s. The estimates of reading achievement from these samples link to six previous reading assessments (1971, 1975, 1979, 1984, 1988, and 1990); the estimates of writing achievement link to three previous writing assessments (1984, 1988, and 1990).
>
> [MS-LTTrend] are age-only samples used for estimating long-term trends in mathematics and science achievement. These samples used the same age definitions and time of testing as were used since 1969 and used the same assessment instruments as were used in previous long-term trend assessments of mathematics and science. As in those previous assessments, the administration of the mathematics and science questions was paced with an audiotape. The estimates of science achievement from these samples link to six previous science

assessments (1969, 1973, 1977, 1982, 1986, and 1990); the estimates of mathematics achievement link to five previous assessments (1973, 1976, 1982, 1986, and 1990).

The Trial State Assessment Samples. These samples are labeled [Math-State] and [Rdg-State] in Table 1-1. 9[Math-State] and 9[Rdg-State] are samples of fourth-grade public-school students from each of the states and jurisdictions participating in the 1992 Trial State Assessment. 13[Math-State] are samples of eighth-grade public school students from the same participating states and jurisdictions. The assessment booklets were the same print-administered booklets as those used for the matching samples 9[Math-MainP], 9[Rdg-MainP], and 13[Math-MainP], but the administrative procedures varied from that of the national assessment in that state personnel collected the data. Students participating in the mathematics assessment were also administered a special set of items measuring estimation skills. These items were presented to students in a separate booklet, accompanied by paced audiotape, and were the same items that were included in the [Math-MainT] assessment.

1.1.2 NAEP Assessments Since 1969

Table 1-2 shows the subject areas, grades, and ages assessed since the NAEP project began in 1969. As can be seen, in addition to the 1992 subject areas of reading, mathematics, science, and writing, several other subject areas have been assessed over the years—social studies, civics, U.S. history, geography, citizenship, literature, music, career development, art, and computer competence. Many subject areas have been reassessed periodically to measure trends over time.

Assessments were conducted annually through 1980, but budget restrictions since then have reduced data collection to a biennial basis. Since its inception, NAEP has assessed 9-year-olds, 13-year-olds, and in-school 17-year-olds, although the age definitions changed in 1986 and again in 1988. Because of budget restrictions, NAEP no longer routinely assesses out-of-school 17-year-olds or young adults. (A separate assessment of young adults of ages 21 to 25 was conducted in 1985 under a separate grant.)

The table also indicates that in 1984, NAEP began gathering data by grade as well as by age, a practice that has been continued in assessments since then. It should be noted that somewhat different age definitions were used in the 1984, 1986, and 1988 assessments. In the 1984 assessment, the younger two ages were defined on a calendar-year basis while the 17-year-olds were defined on an October 1 to September 30 basis. This resulted in modal grades of 4, 8, and 11. To allow for age cohorts that were exactly four years apart, in the 1986 main assessment, all ages were defined on an October 1 to September 30 basis, resulting in modal grades of 3, 7, and 11. Special studies (Kaplan, Beaton, Johnson, & Johnson, 1988) were conducted to measure the effect of the changes in age definition. Because of problems encountered in assessing third graders, in 1988 the ages were redefined on a calendar-year basis, with the modal grades being 4, 8, and 12. These were the age definitions used in the 1990 and 1992 assessments.

Table 1-2
National Assessment of Educational Progress
Subject Areas, Grades, and Ages Assessed: 1969-1992

Assessment Year	Subject Area(s)	Grades/Ages Assessed										
		Grade 3	Grade 4	Age 9	Grade 7	Grade 8	Age 13	Grade 11	Grade 12	Age 17	Age 17OS*	Adult
1969-70	Science			X			X			X	X	X
1970-71	Reading			X			X			X	X	X
	Literature			X			X			X	X	X
1971-72	Music			X			X			X	X	X
	Social Studies			X			X			X	X	X
1972-73	Science			X			X			X	X	X
	Mathematics			X			X			X	X	X
1973-74	Career and Occupational Development			X			X			X	X	X
	Writing			X			X			X	X	
1974-75	Reading			X			X			X	X	
	Art			X			X			X	X	
	Index of Basic Skills									X	X	
1975-76	Citizenship/Social Studies			X			X			X	X	
	Mathematics**						X			X	X	
1976-77	Science			X			X			X		
	Basic Life Skills**									X		
	Science, Reading, Health**											X
1977-78	Mathematics			X			X			X		
	Consumer Skills**									X		
1978-79	Writing, Art, and Music			X			X			X		

---

* Age 17 students who had dropped out of school or had graduated prior to assessment.

** Small, special-interest assessment conducted on limited samples at specific grades or ages.

Table 1-2 (continued)
National Assessment of Educational Progress
Subject Areas, Grades, and Ages Assessed: 1969-1992

Assessment Year	Subject Area(s)	Grades/Ages Assessed										
		Grade 3	Grade 4	Age 9	Grade 7	Grade 8	Age 13	Grade 11	Grade 12	Age 17	Age 17OS*	Adult
1979-80	Reading/Literature			X			X			X	X	
	Art						X					
1981-82	Science**			X			X			X		
	Mathematics and Citizenship/Social Studies			X			X			X		
1984	Reading		X	X		X	X	X		X		
	Writing		X	X		X	X	X		X		
	Reading (long-term trend)		X	X		X	X	X		X		
	Writing (long-term trend)		X	X		X	X	X		X		
1985	Adult Literacy**											X
1986	Reading	X		X	X		X	X		X		
	Mathematics	X		X	X		X	X		X		
	Science	X		X	X		X	X		X		
	Computer Competence	X		X	X		X	X		X		
	U.S. History**							X		X		
	Literature**							X		X		
	Reading (long-term trend)		X	X		X	X	X		X		
	Mathematics (long-term trend)		X	X		X	X	X		X		
	Science (long-term trend)		X	X		X	X	X		X		

---

* Age 17 students who had dropped out of school or had graduated prior to assessment.

** Small, special-interest assessment conducted on limited samples at specific grades or ages.

Table 1-2 (continued)
National Assessment of Educational Progress
Subject Areas, Grades, and Ages Assessed: 1969-1992

Assessment Year	Subject Area(s)	Grades/Ages Assessed										
		Grade 3	Grade 4	Age 9	Grade 7	Grade 8	Age 13	Grade 11	Grade 12	Age 17	Age 17OS*	Adult
1988	Reading		X	X		X	X		X	X		
	Writing		X	X		X	X		X	X		
	Civics		X	X		X	X		X	X		
	U.S. History		X	X		X	X		X	X		
	Document Literacy**					X	X		X	X		
	Geography**								X	X		
	Reading (long-term trend)		X	X		X	X	X		X		
	Writing (long-term trend)		X	X		X	X	X		X		
	Mathematics (long-term trend)			X			X	X		X		
	Science (long-term trend)			X			X	X		X		
	Civics (long-term trend)						X			X		
1990	Reading		X	X		X	X		X	X		
	Mathematics		X	X		X	X		X	X		
	Science		X	X		X	X		X	X		
	Reading (long-term trend)		X	X		X	X	X		X		
	Writing (long-term trend)		X	X		X	X	X		X		
	Mathematics (long-term trend)			X			X			X		
	Science (long-term trend)			X			X			X		
	Trial State Mathematics					X						

---

* Age 17 students who had dropped out of school or had graduated prior to assessment.

** Small, special-interest assessment conducted on limited samples at specific grades or ages.

Table 1-2 (continued)
National Assessment of Educational Progress
Subject Areas, Grades, and Ages Assessed: 1969-1992

Assessment Year	Subject Area(s)	Grades/Ages Assessed										
		Grade 3	**Grade 4**	**Age 9**	**Grade 7**	**Grade 8**	**Age 13**	**Grade 11**	**Grade 12**	**Age 17**	**Age 17OS***	**Adult**
1992	Reading		X	X		X	X		X	X		
	Writing		X	X		X	X		X	X		
	Mathematics		X	X		X	X		X	X		
	Reading (long-term trend)		X	X		X	X	X		X		
	Writing (long-term trend)		X	X		X	X	X		X		
	Mathematics (long-term trend)			X			X			X		
	Science (long-term trend)			X			X			X		
	Trial State Mathematics		X			X						
	Trial State Reading		X									

---

* Age 17 students who had dropped out of school or had graduated prior to assessment.

** Small, special-interest assessment conducted on limited samples at specific grades or ages.

47

46

1.2 DEVELOPMENT OF ASSESSMENT OBJECTIVES, ITEMS, AND BACKGROUND QUESTIONS

In 1992, NAEP conducted main assessments of students at all three age/grade levels in reading, mathematics, and writing. These assessments entailed the generation of a large number of cognitive items. In addition, a large number of background and attitude questions were asked of students and information was collected from principals and teachers. Details on the item development procedures followed for the 1992 main assessment are given in Chapter 2; this section provides an overview. (In addition to the main assessment, long-term trend studies were conducted in reading, mathematics, science, and writing. Since the instruments used for these studies consisted entirely of items used in previous assessments, no developmental tasks were required for their use in the 1992 assessment.)

In addition to the cognitive items, several questionnaires were developed: a common student background questionnaire given to all assessed students of a given age/grade, a school characteristics and policies questionnaire, teacher questionnaires for teachers of fourth- and eighth-grade students, an excluded student questionnaire, and a principal questionnaire. Each of these questionnaires was developed through a broad-based consensus process.

All items in the assessment underwent extensive reviews by subject area and measurement specialists, as well as careful scrutiny to eliminate any potential bias or lack of sensitivity to any group. Further, the items were field tested on a representative group of students. Based on the results of the field test, items were revised or modified as necessary and then again reviewed for bias. With the help of staff and outside reviewers, the Instrument Development Panels selected the items to include in the assessment.

The assessment instruments included multiple-choice, short constructed-response, and extended constructed-response questions in every subject area except writing, which consisted entirely of extended constructed-response questions. The constructed-response questions were professionally scored as described in Chapter 7.

1.3 THE 1992 SAMPLE DESIGN

The sample for the 1992 NAEP assessment was selected using a complex four-stage sample design involving the sampling of students from selected schools within 94 selected geographic areas, called primary sampling units, across the United States. The sample design was similar to that used in 1986, 1988, and 1990 and is described in detail by Westat, Inc., the firm subcontracted by ETS to select the sample, in *1992 National Assessment of Educational Progress Sampling and Weighting Procedures, Final Report* (Wallace & Rust, 1993). The following sections provide an overview of each of the four stages of the sampling design with further details given in Chapter 3.

Stage 1: Primary Sampling Units

In the first stage of sampling, the United States (the 50 states and the District of Columbia) was divided into geographic primary sampling units (PSUs). Each PSU met a minimum size requirement and generally comprised either a metropolitan statistical area (MSA), a single county, or a group of contiguous counties. The PSUs were classified into four regions (Northeast, Southeast, Central, West), each containing about one-fourth of the U.S. population. In each region, PSUs were additionally classified as MSA or nonMSA. In the Southeast and West regions, the PSUs in which 20 percent of the population in the 1980 Census was either Black or Hispanic were further classified as high-minority, while the remaining PSUs in those regions were classified as not high-minority. This resulted in twelve subuniverses of PSUs.

Ninety-four PSUs were selected for the 1992 assessment. Thirty-four PSUs were designated as certainty units because of their size, as it was cost effective to include them in the sample with certainty. Within each major stratum (subuniverse), further stratification was achieved by ordering the noncertainty PSUs according to several additional socioeconomic characteristics. Sixty PSUs were selected, one per stratum from each of the noncertainty strata, with probability proportional to size (the number of school-age children from the 1980 census). To enlarge the samples of Black and Hispanic students, thereby enhancing the reliability of estimates for these groups, PSUs from the high-minority subuniverses were sampled at twice the rate of those from the other subuniverses. This was achieved by creating smaller strata within the high-minority subuniverses. All 94 PSUs were used for both the main assessments and the long-term trend assessments of all three age classes.[2]

Stage 2: Sampling Schools

In the second stage of sampling, the public schools (including Bureau of Indian Affairs schools and Department of Defense schools) and private schools (including Catholic schools) within each of the 94 PSUs were listed according to the three age/grades. An independent sample of schools was selected separately for each of the age/grades so that some schools were selected for assessment of two age/grades, and a few were selected for all three. Schools within each PSU were selected (without replacement) with probabilities proportional to assigned measures of size with oversampling of private schools and of schools with high minority enrollment. Overall probabilities of selection for high-minority schools were twice those for other schools while the probabilities of selection of private schools were triple those of low-minority public schools of the same size. The increased probabilities of selection enlarged the samples of Black and Hispanic students and the samples of students from private schools, thereby enhancing the reliability of estimates for these groups. Details of the probabilities used for school selection appear in Chapter 3.

The samples of schools for the long-term trend assessments were drawn in a manner very similar to that used for the main assessments. The chief difference in the two samples was that private schools and schools with high minority enrollment were not oversampled for these assessments.

---

[2] The term "age class" is used in this report when it is appropriate to discuss one of the three student cohorts in a general way (not necessarily in reference to a specific sample). For the 1992 assessment, age class 9 refers to age 9 and age 9/grade 4 students, age class 13 refers to age 13 and age 13/grade 8 students, and age class 17 includes the age 17, age 17/grade 11, and age 17/grade 12 students.

The overall school cooperation rate exceeded 80 percent at each age/grade. In certain instances, refusing schools were replaced by substitutes according to the rules indicated in Chapter 3.

Stage 3: Assigning Assessment Sessions to Schools

In the third stage of sampling, assessment sessions were assigned to the sampled schools, as described in Chapter 3. An assessment session typically consisted of 25 to 30 students, all of whom could be assessed following the same procedures. There were two general types of sessions in the 1992 assessment: (1) tape sessions, where every student was administered the same booklet and where audiotape prompts paced the students through at least part of the booklet, and (2) print sessions, where a number of distinct booklets were administered and where no audiotape pacing was used. (Print sessions are also called spiral sessions, since the assessment booklets were spiraled for administration—see section 1.4.1.) The assignment of sessions to schools was designed to maximize the number of session types conducted within each PSU, where each session type corresponded to a separate sample of the population of students.

Stage 4: Sampling Students

In the fourth stage of sampling, a consolidated list was prepared for each school of all grade-eligible and age-eligible students for the age class for which the school was selected. To provide the target sample size, a systematic selection of eligible students was made from this list, if necessary. In small and medium-sized schools all eligible students were in the sample. For schools assigned to more than a single session type, students were assigned by Westat district supervisors to one of the various session types using specified procedures. No student was assigned to more than one session.

Stage 4a: Excluded Students

It is NAEP's intention to assess all selected students. However, certain selected students were judged by school authorities as being incapable of participating meaningfully in the assessment. For each of these students, school staff completed an excluded student questionnaire, listing the reason for exclusion and providing some background information.

Specific guidelines for exclusion were provided for all samples in the 1992 assessment. However, somewhat different criteria were used for the long-term trend samples than for the main assessment samples. The exclusion guidelines for the long-term trend samples were the same as those used in previous assessments. Three types of students could be excluded under these guidelines—non-English speaking students, educable mentally retarded students who were judged incapable of meaningfully responding to exercises appropriate to their age level, and students so functionally disabled that they could not perform in the NAEP assessment situation.

The criteria used for the main assessments provided more specific rules for exclusion. A student identified as having Limited English Proficiency (LEP) could be excluded if the student was a native speaker of a language other than English, had been enrolled in an English-speaking school for less than two years, and was judged incapable of taking part in the assessment. Students to be excluded for reasons other than LEP had to be special education students with Individualized Education Plans (IEP), or equivalent designation, who were mainstreamed less than 50 percent of the time in academic subjects and/or were considered unassessable by the IEP team.

Stage 4b: Sampling Teachers

The mathematics teachers of fourth-grade and eighth-grade students participating in the main assessment of mathematics, the reading teachers of fourth-grade students participating in the main assessment of reading, and the writing instructors of eighth-grade students participating in the main assessment of writing were identified and asked to complete a questionnaire (described in Chapter 4) about their background and experiences and about instructional practices, by class, for any classes containing assessed students.

Stage 4c: The School and Principal Questionnaires

A school characteristics and policies questionnaire was mailed to every sampled school by Westat before the assessment. The Westat supervisor then collected the questionnaires and returned them to ETS. The school characteristics and policies questionnaire is described in Chapter 4. The principal questionnaire, distributed to the principal of each sampled school by Westat before the assessment, was used to estimate the number of age/grade-eligible students and to determine the correct "size and type of community" classification for each school.

1.4 ASSESSMENT INSTRUMENTS

Five types of instruments were used in the 1992 assessment: student assessment booklets (which included the student common background questionnaire), excluded student questionnaires, teacher questionnaires, a principal questionnaire, and a school characteristics and policies questionnaire. This section provides an overview of these instruments; more detailed information can be found in Chapter 4.

1.4.1 Student Assessment Booklets—Main Assessment

The student assessment booklets for the 1992 assessment contained both cognitive and noncognitive questions. Each assessment booklet contained common background questions, subject-specific background questions, and between one and three sections of cognitive questions.

As in previous assessments, the assembly of most cognitive items into booklets and their subsequent assignment to assessed students was determined by a *focused balanced incomplete block* (focused-BIB) design with *spiraled* administration. The first step in implementing BIB spiraling is to divide the items within a subject area into units called blocks, where each block is designed to take 15 minutes (for mathematics) or 25 or 50 minutes (for reading and writing) to complete. (The long-term assessments of reading, writing, mathematics, and science used previously created 15-minute blocks of items—the details of the long-term assessment booklets are given later.)

The 15- and 25-minute blocks were assembled into booklets containing the background questions and two (for the 25-minute blocks) or three (for the 15-minute blocks) blocks of subject area items according to a balanced incomplete block design. In a balanced incomplete block design, the subject area blocks are assigned to booklets in such a way that each block appears in the same number of booklets and every pair of blocks of a certain type appears together in at least one booklet. This is the *balanced* part of the method; the *incomplete* part refers to the fact that no booklet contains all items and hence incomplete data is yielded for each assessed student.

As in 1988 and 1990, the BIB design was *focused*—that is, each block of items within each of the subject areas was paired with other blocks within that subject area but not with blocks of items from other subject areas. The focused-BIB design for mathematics was completely balanced in that each block of mathematics items for a given age/grade was paired with every other block of mathematics items for that age/grade. The mathematics design called for 13 15-minute blocks of cognitive items at a given age/grade to be assembled into 26 booklets according to the design shown in Table 1-3. Each block contained both multiple-choice and constructed-response questions. Additionally, extended constructed-response questions were present in many of the blocks. Each booklet contains three blocks each, where each pair of the 13 blocks appeared in exactly one booklet and where each block appeared in six booklets—twice as the first cognitive block in the booklet, twice as the second, and twice as the third. Each booklet also contained five minutes of common background questions and five minutes of mathematics-related background questions.

Besides the focused-BIB booklets, additional booklets were created to measure estimation and complex problem-solving skills. These booklets were administered with an accompanying paced audiotape.

The instruments for the 1992 assessment of reading for age 9/grade 4 included eight 25-minute cognitive blocks of items while the instruments for each of the two older age/grades included nine 25-minute blocks plus two to three 50-minute blocks of items. Each 25-minute block contained one relatively long reading passage and approximately 15 constructed-response and multiple-choice questions. At the two older age/grades, three of the nine 25-minute blocks focused on each of three purposes of reading: Reading for Literary Experience, Reading for Information, and Reading to Perform a Task.

The three literary experience blocks (L_1, L_2, L_3,), the three information blocks (I_1, I_2, I_3), and the three task-oriented blocks (T_1, T_2, T_3) at each of age 13/grade 8 and age 17/grade 12 were paired into 18 booklets according to the design shown in Table 1-4.

This design is partially balanced in the sense that every block within a given purpose of reading is paired with every other block measuring the same purpose but is only paired with one of the three available blocks measuring one of the other two purposes (a fully balanced design in which all 9 blocks are paired together would require an additional 18 booklets). In this design, every block appears in exactly four booklets, twice in booklets measuring a single purpose of

Table 1-3

Main Assessment Mathematics Booklet Design

Booklet	Blocks		
1	1	2	5
2	2	3	6
3	3	4	7
4	4	5	8
5	5	6	9
6	6	7	10
7	7	8	11
8	8	9	12
9	9	10	13
10	10	11	1
11	11	12	2
12	12	13	3
13	13	1	4
14	1	3	8
15	2	4	9
16	3	5	10
17	4	6	11
18	5	7	12
19	6	8	13
20	7	9	1
21	8	10	2
22	9	11	3
23	10	12	4
24	11	13	5
25	12	1	6
26	13	2	7

Table 1-4

Main Assessment Reading Booklet Design
Age 13/Grade 8 and Age 17/Grade 12

Booklet	Blocks
1	L_1 L_2
2	L_2 L_3
3	L_3 L_1
4	I_1 I_2
5	I_2 I_3
6	I_3 I_1
7	T_1 T_2
8	T_2 T_3
9	T_3 T_1
10	L_1 I_2
11	I_3 L_2
12	L_3 I_1
13	I_1 T_2
14	I_2 T_3
15	T_1 I_3
16	L_2 T_1
17	T_2 L_3
18	T_3 T_1

Table 1-5

Main Assessment Reading Booklet Design
Age 9/Grade 4

Booklet	Blocks
1	L_1 L_2
2	L_2 L_3
3	L_3 L_4
4	L_4 L_1
5	L_1 L_3
6	L_2 L_4
7	I_1 I_2
8	I_2 I_3
9	I_3 I_4
10	I_4 I_1
11	I_1 I_3
12	I_2 I_4
13	I_3 L_1
14	I_4 L_2
15	L_3 I_1
16	L_4 I_2

reading and twice in booklets measuring two purposes. The order of administration is balanced, with each block appearing twice as the first cognitive block in the booklet and twice as the second. Furthermore, to partially balance context effects in the booklets measuring two of the purposes of reading, each purpose appears first in at least one booklet and second in at least one other.

At age 9/grade 4, four of the blocks focused on Reading for Literary Experience and four more focused on Reading for Information. For age 9/grade 4, the four literary experience blocks (L_1, L_2, L_3, L_4) and the four information blocks (I_1, L_2, I_3, I_4) were paired into 16 booklets as shown in Table 1-5.

Like the design for the two older age/grades, this design is partially balanced in that every block within a given purpose of reading is paired with every other block measuring the same purpose but is only paired with one block measuring the other of the two purposes. Every block appears in four booklets, three times within booklets measuring the same purpose and once in a booklet measuring both purposes. The order of administration is balanced with each block appearing twice in the first position and twice in the second position of a booklet. Context effects in the four booklets containing both purposes are balanced since each purpose appears first in two booklets and second in the remaining two.

Each booklet at each age/grade also contained two to three minutes of a common core of background questions and seven to eight minutes of general and reading-related background questions. At age 13/grade 8 and age 17/grade 12, additional booklets were created, each consisting of the common and reading-related background questions and a single block requiring 50 minutes for completion. Two booklets were created for age 13/grade 8, one focusing on Reading for Literary Experience and the other focusing on Reading for Information. For age 17/grade 12, three booklets were created, one focusing on Reading for Literary Experience and two focusing on Reading for Information.

The instruments for the 1992 assessment of writing included nine 25-minute blocks of extended constructed-response questions at each age/grade. Each block contained a single prompt requiring students to complete a writing task within a 25-minute period; of the nine 25-minute prompts, three each were for the three purposes of writing (informative, persuasive, and narrative). Additionally, two 50-minute prompts (one narrative and one informative) were administered to the age 13/grade 8 students and three 50-minute prompts (one from each of the three purposes of writing) were administered to the age 17/grade 12 students. The 25-minute prompts were paired into 18 booklets according to the BIB design for reading for the two older age/grades shown in Table 1-4; the 50-minute prompts appeared as the single prompt in a booklet. Each booklet also contained two to three minutes of a common core of common background questions and seven to eight minutes of general and writing-related background questions.

For the main assessment, 62 different booklets were assembled for age 9/grade 4, 67 were assembled for age 13/grade 8, and 69 were assembled for age 17/grade 12. Each booklet consisted of cognitive items from a single subject area. Within each subject area, certain cognitive items were presented to two or three ages.

These booklets were spiraled within subject area and type of administration (paced or print) and placed into bundles. Spiraling involves interleaving the booklets in regular (systematic) sequence so that each booklet appears an appropriate number of times in the sample. Booklets were packaged together in bundles large enough to accommodate a typical assessment session. The bundles were designed so that each booklet would appear equally often in each position in a bundle.

The final step in the BIB-spiraling procedure was the assigning of the booklets to the assessed students. The students within an assessment session were assigned booklets in the order in which the booklets were bundled. As a result, typically, each student in an assessment session received a different booklet and, even in schools with multiple sessions, only a few students received the same booklet or block of items. In the 1992 BIB-spiral design, representative and randomly equivalent national samples of about 2,350 grade- or age-eligible students responded to each item (resulting in samples of about 2,000 students eligible by age and 2,000 eligible by grade).

The BIB-spiral design permits the estimation of correlations between all items within a content area and the estimation of correlations of estimates of proficiency between content areas within a subject area. Furthermore, since the spiral design presents each block of items to fewer persons in any school, but to more schools, than would a simpler matrix sampling design, the cluster effect is markedly reduced, leading to a sample with high statistical efficiency.

A further benefit of the BIB-spiral design is that it balances the position of items across booklets. In a simple matrix sampling scheme, each item appears in only one position in a booklet. In particular, the same items always occur last in a booklet, resulting in potential underestimates of student ability because of fatigue factors. In the BIB-spiral design, each block of items occurs once in each block position, first and second for the two-block designs and first, second, and third for the three-block designs. This means that, of all students administered an item with the three-block design, one-third were presented the item when it appeared in the first part of a booklet, one-third were presented the item when it appeared in the middle of a booklet, and one-third were presented the item when it appeared in the last part of a booklet. (For the two-block BIB design, half of the students administered an item were presented the item in the front half of a booklet; the other half were presented the item in the latter half of a booklet). As a result of this balancing, NAEP has found that in assessments that rely more heavily on reading or on constructed responses, students generally do less well on the last block of questions in a booklet, regardless of the content. However, also as a result of this balancing, NAEP results are not unduly influenced by the fact that students tend to perform less well on items occurring at the end of booklets.

The spiral design does preclude the use of audiotape pacing. Since each student within a session responds to a different set of items, audiotaped administration would be unmanageable. The instructions and the items themselves must be read by the student. For this reason, simple matrix sampling designs, where every student in an assessment session receives the same booklet, are used in the special cases requiring audiotape pacing, such as the assessment of mathematics estimation and complex problem-solving abilities.

1.4.2 Student Assessment Booklets—Long-term Trend Samples

There were two distinct long-term trend samples in the 1992 assessment, each of which required reprinting booklets used in previous assessments:

Reading-Writing Long-term Trend: Six booklets were used at each of the three age/grades for the purposes of measuring long-term trends in reading and writing. These booklets were identical to booklets used in both the 1984 assessments of reading and writing and in the 1988 and 1990 long-term trend assessments of those subjects. Each booklet consisted of a common background block and three cognitive blocks, either two reading and one writing or one reading and two writing. All cognitive blocks also contained subject-related background questions. The booklets were administered without audiotape and were spiraled together for administration.

Mathematics-Science Long-term Trend: These instruments were used for the measurement of mathematics and science and were identical to booklets administered in 1990. These booklets contained 15-minute blocks of mathematics and science items; each mathematics block and each science block was administered using audiotape pacing. (At the younger two ages, the booklets also contain a block of reading items, which was print-administered.) There were three booklets each at age 9 and age 13 and two booklets at age 17. Combined, the booklets at an age contain three blocks of mathematics items and three blocks of science items. Because of the audiotape pacing, each booklet was administered in a separate session.

1.4.3 Other Instruments

Besides the student assessment booklets, other instruments provided data relating to the assessment:

Excluded student questionnaires. Some students that were selected for the sample may have been deemed unassessable by school personnel. In every such case, school personnel were asked to fill out an Excluded Student Questionnaire that included questions about the reason that the student was excluded and some questions about the student's background.

Teacher questionnaires were administered to the teachers of all students participating in the fourth-grade reading, the eighth-grade writing, and the fourth- and eighth-grade mathematics assessments. The teacher questionnaire included a general section that contained questions about the teacher's background and experience. The rest of the questionnaire contained questions about instructional practices, by class, for any classes containing assessed students.

School characteristics and policies questionnaires were completed by school principals or their representatives, who provided information about school administration, staffing patterns, special programs, subject requirements, and school resources.

Principal questionnaires, distributed to the principals of sampled schools before the assessment, were used to estimate the number of age/grade-eligible students and to determine the correct "size and type of community" classification for each school.

1.5 FIELD OPERATIONS AND DATA COLLECTION

Field operations and data collection for the 1992 assessment were the responsibility of Westat, Inc., and are documented in Chapter 5. The field operation was conducted by a staff at Westat's home office and a larger staff in the field. The Westat home office staff coordinated all activities related to field operations and managed materials distribution and home-office receipt of assessment reporting forms. The field staff consisted of area supervisor, assessment supervisors, and exercise administrators. The assessment supervisors, who were trained by Westat, were each responsible for the assessment activities in one or more PSUs. Although ETS made initial contact with participating school districts, each assessment supervisor was primarily responsible for making follow-up contacts with these districts, recruiting and training exercise administrators to work with them in administering the assessment sessions, arranging the assessment sessions, and selecting the sample of students to be assessed within each school. The assessment supervisors and the exercise administrators administered the assessments, filled out the necessary forms, performed process control, and shipped the assessment booklets and forms to National Computer Systems (NCS), the subcontractor responsible for processing NAEP materials and data.

Gaining school cooperation was the joint responsibility of Westat and ETS staff. ETS made the preliminary contacts preparatory to obtaining school cooperation by first contacting the Chief State School Officers, informing them that schools within their states had been selected for the assessment and, in a later letter, listing the selected schools and districts. Later mailings were sent to superintendents of public schools and parochial schools and principals of private schools for all schools selected in the assessment. These materials provided an explanation of NAEP, a list of the selected schools in the official's jurisdiction, and a cover letter explaining that a Westat district supervisor would contact them to set up an introductory meeting. Westat district supervisors then scheduled and conducted introductory meetings, worked with the schools to schedule the assessments, and, with the exercise administrators, conducted the assessments. The overall cooperation rate of schools originally selected was 85 percent for the 1992 main samples and 84 percent for the long-term trend samples. Further detail on school participation rates is given in sections 3.2 and 3.3 of Chapter 3.

The main assessment sessions were conducted between January 6 and April 3, 1992 at all age/grade levels. The age 9/grade 4 long-term trend assessments were carried out between January 6 and March 13; the age 17/grade 11 long-term trend samples were conducted between March 16 and May 15, 1992. The age 13/grade 8 long-term trend assessments were carried out between October 7 and December 13, 1991.

An automated management system tracked and recorded the progress of field work throughout the 1992 assessment period. In addition, progress was constantly monitored through telephone reports held between the area supervisors and the assessment supervisors and between the area supervisors and the home office staff.

Both Westat and ETS participated in the quality control of the field administration, which involved on-site visits by Westat and ETS staff to verify the sampling of the students and to observe the conduct of the assessment by the supervisors and the exercise administrators.

1.6 MATERIALS AND DATA PROCESSING

After completing an assessment session, Westat field supervisors and exercise administrators shipped the assessment booklets and forms from the field to National Computer Systems for entry into computer files, professional scoring, and creating the data files for transmittal to ETS. Careful checking assured that all data from the field were received. More than 215,000 booklets and questionnaires were received and processed for the national portion of the 1992 assessment. The extensive processing of these data is detailed in Chapter 6.

The student data were transcribed into machine-readable form by scanning the student instruments with an optical scanning machine. An intelligent data entry system was used for resolution of the scanned data, the entry of documents rejected by the scanning machine, and the entry of information from the questionnaires. Additionally, each piece of input data was checked to verify that it was of an acceptable type, that it was within a specified range or ranges of values, and that it was consistent with other data values. The entry and editing of materials is discussed in Chapter 6.

1.7 PROFESSIONAL SCORING

Items requiring a written response from the student (constructed-response items) were included in the main assessments in reading, mathematics, and writing, the Trial State Assessment in mathematics, and the long-term trend assessments in reading, mathematics, science, and writing. More than 2.3 million constructed responses were read and marked by the professional scoring staff for the national portion of the 1992 assessment.

Chapter 7 describes the professional scoring operation, including an overview of the scoring guides, the training procedures, and the scoring process for each subject area.

1.8 CREATION OF THE DATABASE

Before any analyses could begin, the student response data, school, teacher, and excluded student questionnaire data, and all sampling weights had to be integrated into a coherent and comprehensive database. This database was used for all analyses. The database was also the source for the creation of three NAEP database products—the integrated information database, the restricted-used data files, and the secondary-use data files. The quality of the data resulting from the complete data entry system, from the actual instruments collected in the field to the final machine-readable database used in analysis, was verified by selecting field instruments at random and performing a character-by-character comparison of these instruments with their representations in the final database. Chapter 8 provides details on the database, quality control activities, and database products.

Chapter 2

DEVELOPING THE NAEP OBJECTIVES, ITEMS, AND BACKGROUND QUESTIONS FOR THE 1992 ASSESSMENTS OF READING, MATHEMATICS, AND WRITING

Jay R. Campbell, Stephen Lazer, and Ina V.S. Mullis

Educational Testing Service

The subject areas constituting the main 1992 assessments were specified by the 1988 legislation authorizing the National Assessment of Educational Progress and included reading, mathematics, and writing.[1] Long-term trend assessments were conducted in these subjects as well as in science; these assessments used items that were developed as part of previous assessments. Additional data came from the Trial State Assessment Program, which in 1992 assessed mathematics at grades 4 and 8 and reading at grade 4 in representative samples of public-school students in 41 participating states, the District of Columbia, and two territories. A summary of each main assessment subject area follows:

Reading: For the national assessment, a newly developed reading assessment was administered at grades 4, 8, and 12. This assessment was designed around questions requiring in-depth analysis of authentic, naturally occurring reading materials. A mixture of multiple-choice, short constructed-response, and extended constructed-response questions made up the survey; in aggregate well over half of the student assessment time was spent answering constructed-response rather than multiple-choice questions. In addition, a special study entitled the Integrated Reading Performance Record (IRPR) was conducted at grade 4. In this study, a nationally representative subset of fourth graders who had participated in the main reading assessment read aloud and answered questions about their reading habits and attitudes on audiotape. For the participants in the Trial State Assessment Program, the grade 4 assessment was administered without the special study.

Mathematics: For the nation, the assessment that had been developed in 1990 was nearly doubled in scope for 1992 and administered at grades 4, 8, and 12. Assessment tasks included the use of four-function calculators at grade 4, scientific calculators at grades 8 and 12, and open-ended problem-solving questions at all grades. Manipulatives, rulers, and protractors also were available for use with portions of the assessment. Extended constructed-response questions were used on a wide scale for the first time in a NAEP mathematics assessment. In addition, estimation was assessed using audiotapes that paced students through the questions, and complex problem-solving skills were assessed for the nation in special study blocks. For the participating states and jurisdictions in the Trial State Assessment, the mathematics assessment was administered at grades 4 and 8 which provided trend information at grade 8.

---

[1] Copies of the frameworks for these assessments are available from the National Assessment Governing Board.

Writing: A newly developed writing assessment at grades 4, 8, and 12, based entirely on extended constructed-response tasks, was administered to nationally representative samples of students. Participants responded either to two 25-minute prompts or (at grades 8 and 12) to one 50-minute prompt. In accordance with NAEP's 1992 Writing Assessment Framework, students were required to complete informative, narrative, or persuasive writing tasks. The writing assessment also contained a classroom-based study in portfolio assessment at grades 4 and 8.

From its inception, NAEP has developed assessments through a consensus process and the 1992 assessment was no exception. Under the direction of the National Assessment Governing Board (NAGB), educators, scholars, and citizens representative of many diverse constituencies and points of view designed an assessment framework for each of the three subject areas, proposing goals they felt students should achieve in the course of their education. Staff at Educational Testing Service who were subject-area experts in their respective fields worked with subject-area consultants well versed in assessment methodology to develop assessment questions appropriate to the objectives. All questions underwent extensive reviews by subject-matter specialists, measurement specialists, and ETS employees trained to conduct "sensitivity" reviews designed to prevent the inclusion in assessments of materials that might prove offensive or unfair to subpopulations in American society. Questions were assembled and printed into booklets suitable for matrix sampling and then administered either by a trained field staff (for the national program) or by state or local school district staff (for the Trial State Assessment Program) to stratified, multistage probability samples of students.

All 1992 development efforts were governed by four major considerations:

1) As specified in the legislation, the objectives were to be developed through a consensus process involving subject-matter experts, school administrators, teachers, and parents, and the items were to be reviewed carefully for potential bias.

2) As outlined in the ETS proposal for the administration of the NAEP contract, the development of questions for each subject area was to be guided by an Instrument Development Committee that contained considerable overlap with the panelists involved in guiding the consensus process.

3) As described in the *ETS Standards for Quality and Fairness* (Educational Testing Service, 1987), all materials developed at ETS were to be in compliance with specified procedures. Specifically, all questions were to be carefully reviewed for content accuracy, testworthiness, and potential bias.

4) As required by federal regulations, all NAEP items were to be submitted to a complex clearance process. This process involved review of all cognitive items by NCES and NAGB, and review of all background questions by the Office of

Management and Budget (OMB), the Information Management and Compliance Division (IMCD) of the Department of Education, and NCES.

The development effort for the 1992 assessment included questionnaires for students, teachers, and school administrators, in addition to a substantial number of cognitive items for each of the three subject areas.

The following sections include general overviews about setting objectives and developing items and specific details about developing subject-specific objectives and assessments. A list of the consultants who participated in the 1992 development process is included in Appendix A.

2.1 GENERAL OVERVIEW OF THE 1992 ASSESSMENT FRAMEWORKS

The subject-area objectives for each NAEP assessment are determined through a legislatively mandated consensus process. These objectives typically take the form of frameworks or matrices delineating the important content and process areas to be assessed. The various frameworks for the 1992 assessments are described below and discussed in detail by Mullis (1990).

The 1992 **reading** assessment framework is a four-by-three matrix specifying four reading stances and three reading purposes. The stances are initial understanding, developing an interpretation, personal reflection and response, and demonstrating a critical stance. The three global reading purposes are reading for literary experience, reading to gain information, and reading to perform a task. The 1992 reading assessment measured students' abilities to read based on a variety of passages, ranging from informational materials, documents, and news articles to poems, essays, and stories. The stimuli used in NAEP represent the types of materials that students commonly encounter in and out of school and are expected to be capable of reading.

The reading framework was developed under the auspices of the Council of Chief State School Officers under contract to the National Assessment Governing Board. The project involved widespread participation and review by many groups, including a planning committee composed of reading educators, and a steering committee representing policymakers, curriculum developers, and other consumers of NAEP reports. In addition, the framework was reviewed extensively by experts in the field of reading, by state education officials, and by representatives of professional associations. The reading objectives were also the subject of testimony at public hearings arranged to allow for the widest possible participation in the consensus process. While objectives resulting from such a consensus process reflect neither a narrowly defined theoretical framework nor every view of every participant, they do represent the thinking of a broad cross-section of individuals who are expert in the areas of literacy research and reading instruction and who are deeply committed to the improvement of reading in our schools.

NAEP's 1992 **mathematics** assessment was used to measure short-term trends from 1990. The mathematics framework is a five-by-three matrix specifying five content areas (numbers and operations; measurement; geometry; data analysis, statistics, and probability; and algebra and functions) and three process or ability areas (conceptual understanding, procedural knowledge, and problem solving).

The mathematics objectives were developed under the auspices of the Council of Chief State School Officers through a special NAEP Planning Project sponsored by the National Center for Education Statistics and the National Science Foundation. This project involved widespread participation and review by many groups, including an objectives committee of mathematics educators; a steering committee of 18 members representing policy makers, practitioners, and citizens at large; distribution to the mathematics supervisor in each state education agency for review by state committees; and reviews by mathematics scholars, the National Center for Education Statistics, and NAEP's governing board.

The 1992 **writing** framework focuses on students' abilities to write effectively for three purposes—informative, narrative, and persuasive. The framework also emphasizes students' abilities to manage the writing process and to meet standards of organization, elaboration, and convention.

The writing framework was developed under a consensus process managed by NAGB and under the immediate direction of a writing consensus process committee made up of writing and assessment experts and writing practitioners (see Appendix A). The objectives were reviewed by state education personnel, representatives of business and industry, and writing and assessment experts.

2.2 GENERAL OVERVIEW OF PROCEDURES FOR DEVELOPING THE ITEMS

A carefully developed and tested series of steps, similar to those for past NAEP assessments, was used to create assessment items that reflected the mathematics, reading, and writing objectives and that measured achievement related to them (see sections 2.3—2.5 for more detail). The steps were as follows:

1) Item specifications and prototype items were provided in conjunction with the 1992 frameworks in each subject area.

2) The Instrument Development Committees in each subject area provided guidance to NAEP staff about ways in which the objectives could be measured given the realistic constraints of resources and measurement technology. The committees made recommendations about priorities for the assessment (within the context of the assessment frameworks) and the types of questions to be developed.

3) In mathematics, the existing pool of items to be used to measure change from previous assessments (trend items) was reviewed in detail and trend items were selected.

4) Specialists with subject-matter expertise and skills and experience in creating items according to specifications were identified from both inside and outside ETS to develop and review the assessment questions. The development of items and related materials (passages, stimuli, etc.) therefore involved the input of practitioners from around the country as well as from members of the Instrument Development Committee.

5) Newly created questions were reviewed and revised by staff and external reviewers, including the Instrument Development Committee. The items for the fourth- and eighth-grade mathematics assessments and the fourth-grade reading assessment were also reviewed by representatives from the State Education Agencies. For the reading assessment, blocks of materials and questions were sent to teachers across the country for review.

6) Further language editing and sensitivity reviews were conducted as required by the *ETS Standards For Quality and Fairness*.

7) Field test materials were prepared, including the materials necessary to secure Office of Management and Budget clearance. The clearance packages were also reviewed by representatives of the National Center for Education Statistics and the National Assessment Governing Board had responsibility for approving the cognitive materials.

8) Field tests for the national program were conducted with representative groups of students from across the country.

9) Field-test booklets were scored and the results analyzed.

10) Based on these analyses and the results of the field test, questions were revised or modified and re-edited. They once again went through the required ETS sensitivity review.

11) With the help of staff and outside reviewers, each Instrument Development Committee selected the items or, in the case of reading and writing, the blocks to include in the operational assessments.

12) In mathematics, where blocks do not always move from field test to operational assessment in intact form, items were assembled into blocks with attention given to balancing content coverage and difficulty levels.

13) Each block underwent final content, editorial, and sensitivity reviews.

14) After a final check to ensure that each assessment booklet and each block therein met the overall guidelines for the assessment, the materials were sent for final Office of Management and Budget clearance, National Center for Education Statistics review, and National Assessment Governing Board approval for cognitive materials.

15) The booklets were typeset and printed.

The following sections describe the development of the reading, mathematics, and writing assessments in more detail.

2.3 DEVELOPING THE READING ASSESSMENT

2.3.1 Overview

Similar to all previous NAEP assessments, the framework for the 1992 reading assessment was developed through a broad-based consensus process. To prepare the framework and objectives for the 1992 reading assessment, the National Assessment Governing Board contracted with the Council of Chief State School Officers. The development process involved a steering committee, a planning committee, and Council of Chief State School Officers project staff. Educators, scholars, and citizens, representative of many diverse constituencies and points of view, participated in the national consensus process to design objectives for the reading assessment.

The development of the assessment followed procedures used in a variety of NAEP development projects at ETS. In addition to extensive review by ETS content, sensitivity, and editorial specialists, the assessment questions were reviewed by a wide variety of external experts, including the educators and scholars on the Instrument Development Committee, and teachers around the country. All materials in the operational assessment were systematically field tested.

2.3.2 Development of the Assessment Framework

The National Assessment Governing Board is responsible for guiding NAEP, including the development of the reading assessment objectives and test specifications. Appointed by the Secretary of Education from lists of nominees proposed by the Board itself in various statutory categories, the 24-member board is composed of state, local, and federal officials, as well as educators and members of the public.

The National Assessment Governing Board began the development process for the 1992 reading objectives by conducting a widespread mail review of the objectives for the 1990 reading assessment and by holding a series of public hearings throughout the country. The contract for managing the remainder of the consensus process was awarded to the Council of Chief State School Officers. The development process included the following activities:

- A Steering Committee consisting of members recommended by each of 15 national organizations (see Appendix A) was established to provide guidance for the consensus process. The committee responded to the progress of the project and offered advice. Drafts of each version of the document were sent to members of the committee for review and reaction.

- A Planning Committee (see Appendix A) was established to identify the objectives to be assessed in reading in 1992 and prepare the framework document. The members of this committee consisted of experts in reading, including college professors, an academic dean, a classroom teacher, a school administrator, state level assessment and reading specialists, and a representative of the business community. This committee met with the Steering Committee and as a separate group. A subgroup also met to develop item specifications. Between meetings, members of the committee provided information and reactions to drafts of the framework.

- The project staff at the Council of Chief State School Officers met regularly with staff from the National Assessment Governing Board and the National Center for Education Statistics to discuss progress made by the Steering and Planning committees.

During this development process, input and reactions were continually sought from a wide range of members of the reading field, experts in assessment, school administrators, and state staff in reading assessment. In particular, the process was informed by innovative state assessment efforts and work being done by the Center for the Learning and Teaching of Literature (Langer, 1989, 1990).

2.3.3 Framework and Assessment Design Principles

The reading objectives framework was designed to focus on reading processes and outcomes, rather than reflect a particular instructional or theoretical approach. It was stated that the framework should focus not on the specific reading skills that lead to outcomes, but rather on the quality of the outcomes themselves. The framework was intended to embody a broad view of reading by addressing the increasing level of literacy needed for employability, personal development, and citizenship. The framework also specified a reliance on contemporary reading research and the use of nontraditional assessment formats that more closely resemble desired classroom activities.

The objectives development was guided by the consideration that the assessment should reflect many of the curricular emphases and objectives in various states, localities, and school districts in addition to what various scholars, practitioners, and interested citizens believed should be included in the curriculum. Accordingly, the committee gave attention to several frames of reference:

- The purpose of the NAEP reading assessment is to provide information about the progress and achievement of students in general rather than to test individual students' ability. NAEP is designed to inform policymakers and the public about reading ability in the United States. Furthermore, NAEP state data can be used to inform states of their students' relative strengths and weaknesses.

- The term "reading literacy" should be used in the broad sense of knowing when to read, how to read, and how to reflect on what has been read. It represents a complex, interactive process that goes beyond basic or functional literacy.

- The reading assessment should use valid and authentic tasks that are both broad and complete in their coverage of important reading behaviors so that the test will be useful and valid, and will demonstrate a close link to desired classroom instruction.

- Every effort should be made to make the best use of available methodology and resources in driving assessment capabilities forward. New types of items and new methods of analysis were recommended for the 1992 NAEP in reading.

- Every effort must be made in developing the assessment to represent a variety of opinions, perspectives, and emphases among professionals, as well as state and local school districts.

2.3.4 Framework for the 1992 Assessment

The framework adopted for the 1992 reading assessment is organized according to a four-by-three matrix of reading *stances* by reading *purposes*. The stances include

- Initial Understanding,
- Developing an Interpretation,
- Personal Reflection and Response, and
- Demonstrating a Critical Stance.

These stances were assessed across three global purposes defined as

- Reading for Literary Experience,
- Reading to Gain Information, and
- Reading to Perform a Task.

Different types of texts were used to assess the various purposes for reading. Students' reading abilities were evaluated in terms of a single purpose for each type of text. At grade 4 only Reading for Literary Experience and Reading to Gain Information were assessed, while all three global purposes were assessed at grades 8 and 12. Figure 2-1 and 2-2 describe the four reading stances and three reading purposes that guided the development of NAEP's 1992 reading assessment, including the Trial State Assessment at grade 4.

For 1992, the Planning Committee was interested in creating an assessment that would be forward-thinking and reflect quality instruction. In recognition that the demands made of readers change as they mature and move through school, it was recommended that the proportion of items had some relation to reading purpose (to perform a task, for literary experience, to gain information). The distribution of items by reading purpose across grade levels recommended in the assessment framework is provided in Table 2-1.

Figure 2-1
Description of Reading Stances

Readers interact with text in various ways as they use background knowledge and understanding of text to construct, extend, and examine meaning. The NAEP reading assessment framework specified four reading stances to be assessed that represent various interactions between readers and texts. These stances are not meant to describe a hierarchy of skills or abilities. Rather, they are intended to describe behaviors that readers at all developmental levels should exhibit.

Initial Understanding

Initial understanding requires a broad, preliminary construction of an understanding of the text. Questions testing this aspect ask the reader to provide an initial impression or unreflected understanding of what was read. In the 1992 NAEP reading assessment, the first question following a passage was usually one testing initial understanding.

Developing an Interpretation

Developing an interpretation requires the reader to go beyond the initial impression to develop a more complete understanding of what was read. Questions testing this aspect require a more specific understanding of the text and involve linking information across parts of the text as well as focusing on specific information.

Personal Reflection and Response

Personal response requires the reader to connect knowledge from the text more extensively with his or her own personal background knowledge and experience. The focus is on how the text relates to personal experience; questions on this aspect ask the readers to reflect and respond from a personal perspective. For the 1992 NAEP reading assessment, personal response questions were typically formatted as constructed-response items to allow for individual possibilities and varied responses.

Demonstrating a Critical Stance

Demonstrating a critical stance requires the reader to stand apart from the text, consider it, and judge it objectively. Questions on this aspect require the reader to perform a variety of tasks such as critical evaluation, comparing and contrasting, application to practical tasks, and understanding the impact of such text features as irony, humor, and organization. These questions focus on the reader as interpreter/critic and require reflection and judgments.

Figure 2-2
Description of Purposes for Reading

Reading involves an interaction between a specific type of text or written material and a reader, who typically has a purpose for reading that is related to the type of text and the context of the reading situation. The 1992 NAEP reading assessment presented three types of text to students representing each of three reading purposes: literary text for literary experience, informational text to gain information, and documents to perform a task. Students' reading skills were evaluated in terms of a single purpose for each type of text.

Reading for Literary Experience

Reading for literary experience involves reading literary text to explore the human condition, to relate narrative events with personal experiences, and to consider the interplay in the selection among emotions, events, and possibilities. Students in the NAEP reading assessment were provided with a wide variety of literary text, such as short stories, poems, fables, historical fiction, science fiction, and mysteries.

Reading to Gain Information

Reading to gain information involves reading informative passages in order to obtain some general or specific information. This often requires a more utilitarian approach to reading that requires the use of certain reading/thinking strategies different from those used for other purposes. In addition, reading to gain information often involves reading and interpreting adjunct aids such as charts, graphs, maps, and tables that provide supplemental or tangential data. Informational passages in the NAEP reading assessment included biographies, science articles, encyclopedia entries, primary and secondary historical accounts, and newspaper editorials.

Reading to Perform a Task

Reading to perform a task involves reading various types of materials for the purpose of applying the information or directions in completing a specific task. The reader's purpose for gaining meaning extends beyond understanding the text to include the accomplishment of a certain activity. Documents requiring students in the NAEP reading assessment to perform a task included directions for creating a time capsule, a bus schedule, a tax form, and instructions on how to write a letter to a senator. In 1992, reading to perform a task was assessed only at grades 8 and 12.

Table 2-1
Percentage Distribution of Items
by Grade and Reading Purpose

Grade	Purposes for Reading		
	Literary Experience	To Gain Information	To Perform a Task
4	55%	45%	(No Scale)
8	40%	40%	20%
12	35%	45%	20%

Readers use a range of cognitive abilities and assume various stances that should be assessed within each of the reading purposes. While reading, students form an initial understanding of the text and connect ideas within the text to generate interpretations. In addition, they extend and elaborate their understanding by responding to the text personally and critically and by relating ideas in the text to prior experiences or knowledge. Table 2-2 shows the distribution of items by reading stance, as specified in the reading framework, for all three grade levels.

Table 2-2
Percentage Distribution of Items
by Reading Stance for Grades 4, 8, and 12

Initial Understanding/ Developing an Interpretation	Personal Response	Critical Stance
33%	33%	33%

2.3.5 Developing the Cognitive Items

The development of cognitive items began with a careful selection of grade-appropriate passages for the assessment. Passages were selected from a pool of reading selections contributed by teachers from across the country. The framework stated that the assessment passages should represent authentic, naturally occurring reading material that students may encounter in and out of school. Furthermore, these passages were to be reproduced in test booklets as they had appeared in their original publications. In some cases materials (such as bus schedules) were provided to students separate from the printed assessment booklet. Final passage selections were made by the Reading Instrument Development Committee. Finally, in order to guide the development of items, passages were outlined or mapped to identify essential elements of the text.

The assessment included constructed-response (short and extended) and multiple-choice items. The decision to use a specific item type was based on a consideration of the most appropriate format for assessing the particular objective. Both types of constructed-response items were designed to provide an in-depth view of students' ability to read thoughtfully and generate their own responses to reading. Short constructed-response questions were used when students needed to respond in only one or two sentences in order to demonstrate full comprehension. Extended constructed-response questions were used when the task required more thoughtful consideration of the text and engagement in more complex reading processes. Multiple-choice items were used when a straightforward, single correct answer was all that was required. Guided by the NAEP reading framework, the Instrument Development Committee monitored the development of all three types of items to assess objectives in the framework.

A carefully developed and proven series of steps was used to create the assessment items. These steps are described earlier in this chapter under section 2.2.

The assessment included different 25-minute and 50-minute "blocks," each consisting of one or more passages and a set of multiple-choice and constructed-response items to assess students' comprehension of the written material. Students were asked to respond to either two 25-minute blocks or one 50-minute block (at grades 8 and 12). The grade 4 assessment included eight 25-minute blocks (four blocks measuring each of the two global purposes for reading assessed at this grade). The instruments at grades 8 and 12 each included nine 25-minute blocks (three blocks measuring each of the global purposes for reading). In addition, the grade 8 assessment included two 50-minute blocks and the grade 12 assessment included three 50-minute blocks. One of the 50-minute blocks at each grade level represented the *NAEP Reader*, in which students were asked to choose a story to read and to answer questions about that story.

The Integrated Reading Performance Record (IRPR) special study was also conducted at grade 4. A special development committee was organized to oversee the development of the IRPR. This committee had membership that overlapped with that of the Reading Instrument Development Committee.

2.3.6 Distribution of Assessment Items

Table 2-3 lists the total number of questions at each grade level in the 1992 assessment. A total of 126 unique multiple-choice items and 171 constructed-response questions make up the 1992 reading assessment. The total number of questions is thus 297. Many of these items are used at more than one grade level. As a result, the sum of the items that appear at each grade level is greater than the total number of unique items.

In the development, every effort was made to meet specified targets. Table 2-4 shows the approximate percentage of aggregate assessment time devoted to each purpose of reading, at each grade level. Actual percentages are based on the classifications agreed upon by NAEP's 1992 Instrument Development Committee.

Table 2-3

Distribution of Questions for the 1992 Reading Assessment

<table>
<tr><th></th><th>Grade 4</th><th>Grade 8</th><th>Grade 12</th></tr>
<tr><td>Grade 4 only</td><td>31 multiple-choice
33 constructed-response</td><td></td><td></td></tr>
<tr><td>Grades 4 & 8 overlap</td><td colspan="2">11 multiple-choice
10 constructed-response</td><td></td></tr>
<tr><td>Grade 8 only</td><td></td><td>25 multiple-choice
42 constructed-response</td><td></td></tr>
<tr><td>Grade 8 & 12 overlap</td><td></td><td colspan="2">20 multiple-choice
27 constructed-response</td></tr>
<tr><td>Grade 12 only</td><td></td><td></td><td>39 multiple-choice
59 constructed-response</td></tr>
<tr><td>Total</td><td>42 multiple-choice
43 constructed-response</td><td>56 multiple-choice
42 constructed-response</td><td>59 multiple-choice
86 constructed-response</td></tr>
</table>

Table 2-4

Target and Actual Percentage Distribution of Questions by Grade and Reading Purpose

	Grade 4		Grade 8		Grade 12	
Reading Purpose	**Target**	**Actual**	**Target**	**Actual**	**Target**	**Actual**
Literary	55%	50%	40%	36%	35%	33%
Informational	45%	50%	40%	36%	45%	42%
Perform a Task	N/A	N/A	20%	28%	20%	25%

Table 2-5 shows the approximate percentage of assessment time devoted to each reading stance. Unlike the purposes for reading, in which individual students did not receive questions in all areas, every student completed tasks involving each of the reading stances. It is recognized that making discrete classifications is difficult for these categories and that independent efforts to classify NAEP questions have led to different results (National Academy of Education, 1992). Also, it has been found that developing personal response questions that are considered equitable across students' different backgrounds and experiences is difficult.

Table 2-5
Target and Actual Percentage Distribution of Questions by Grade and Reading Ability

	Grade 4		**Grade 8**		**Grade 12**	
Reading Stance	**Target**	**Actual**	**Target**	**Actual**	**Target**	**Actual**
Initial Understanding/ Developing an Interpretation	33%	39%	33%	44%	33%	39%
Personal Response	33%	27%	33%	22%	33%	23%
Critical Stance	33%	34%	33%	34%	33%	38%

2.3.7 Reading Background Questionnaires

Research indicates that school, home, and attitudinal variables affect students' reading comprehension and literacy. Therefore, in addition to assessing how well students read, it is important to understand the instructional context in which reading takes place, students' home support for literacy, and their reading habits and attitudes. To gather contextual information, NAEP assessments include background questions designed to provide insight into the factors that may influence reading proficiency in the literary, informational, and document categories assessed.

NAEP includes both general background questionnaires given to participants in all subjects and subject-specific questionnaires for both students and their teachers. The development of the general background questionnaires is discussed below. However, it is worth noting that members of the Reading Instrument Development Committee were consulted on the appropriateness of the issues addressed in all questionnaires that may relate to reading instruction and achievement. Like the cognitive items all background questions were submitted to extensive review and field testing. Recognizing the validity problems inherent in self-reported data, particular attention was given to developing questions that were meaningful and unambiguous and that would encourage accurate reporting.

In addition to the cognitive questions, the 1992 assessment included two five-minute sets of general and reading background questions designed to gather contextual information about students, their instructional and recreational experiences in reading, and their attitudes toward reading. A one-minute questionnaire was given to students at the end of each booklet to determine students' motivation in completing the assessment and their familiarity with assessment tasks.

Reading Student Questionnaires

Three categories of information were represented in the five-minute sections of reading background questions called the **student reading questionnaire** (14 questions at grade 4, 24 questions at grades 8 and 12):

> *Time Spent Studying Reading*: Time spent on task and reading coursework has been shown to be strongly related to reading achievement (Anderson, Hiebert, Scott, & Wilkinson, 1984). Students were asked to describe both the amount of instruction they received in reading and the time spent on reading homework.
>
> *Instructional Practices*: The nature of students' reading instruction is also thought to be related to achievement (Dole, Duffy, Roehler, & Pearson, 1991). Students were asked to report their instructional experiences related to reading in the classroom, including group work, special projects, and writing in response to reading. In addition, they were asked about the instructional practices of their reading teachers and the extent to which the students themselves discussed what they read in class and demonstrated use of skills and strategies.
>
> *Attitudes Towards Reading*: Students' enjoyment of and confidence in their abilities in reading and their perceptions of the usefulness of reading to their present and future lives appear to be related to reading achievement (Guthrie & Greaney, 1991). Students were asked a series of questions about their attitudes and perceptions about reading, such as whether they enjoyed reading and whether they were good in reading.

Teacher Questionnaire

To supplement the information on instruction reported by students, the reading teachers of the fourth graders participating in the NAEP reading assessment were asked to complete a questionnaire about their instructional practices, teaching backgrounds, and characteristics. The teacher questionnaire contained two parts. The first part pertained to the teachers' background and general training. The second part pertained to specific training in teaching reading and the procedures the teacher used for *each class* containing an assessed student.

The Teacher Questionnaire, Part I: Background and General Training (23 questions) included questions pertaining to gender, race/ethnicity, years of teaching experience, certification, degrees, major and minor fields of study, coursework in education, coursework in specific subject areas, amount of in-service training, extent of control over instructional issues, and availability of resources for their classroom. This component of the questionnaire was completed by teachers whose students participated in any subject assessed in NAEP.

The **Teacher Questionnaire, Part II: Training in Reading and Classroom Instructional Information** (56 questions) included questions on the teacher's exposure to various issues related to reading and teaching reading through pre- and in-service training, ability level of students in the class, whether students were assigned to the class by ability level, time on task, homework assignments, frequency of instructional activities used in class, methods of assessing student progress in reading, instructional emphasis given to the reading abilities covered in the assessment, and use of particular resources. This section of the questionnaire was completed only by teachers whose students took part in the reading assessment.

2.3.8 Development of Final Forms

Reading field tests were conducted in February and March 1991 and involved national samples of fourth-, eighth-, and twelfth grades and 9-, 13-, and 17 year olds. The intent of the field test was to try out the items and procedures and to give the contractors (and at grade 4, the states) practice and experience with the proposed materials and procedures. At grade 4, 149 items were field tested (yielding 85 for the operational assessment); the grade 8 field test included 212 items (98 were used operationally); and the grade 12 field test included 216 items (145 were included in the final assessment). These figures may tend to understate the field test effort, since items for the NAEP Reader were counted only once but were analyzed and scored separately depending on students' choices of story. About 500 responses were obtained to each item in the field test.

The field test data were collected, scored, and analyzed in preparation for meetings with the Reading Instrument Development Committee. Using item analysis, which provided the mean percentage of correct responses, the biserial correlations, and the difficulty level for each item in the field test, committee members, ETS test development staff, and NAEP/ETS staff reviewed the materials. In addition, another meeting of representatives from state education agencies was convened to review the field-test results at grade 4. Four objectives guided these reviews: to determine which items were most related to overall student achievement; to determine the need for revisions of items that lacked clarity, or had ineffective item formats; to prioritize items to be included in the assessment; and to determine appropriate timing for assessment items.

Once the committees had selected the items, all items were rechecked for content, measurement, and sensitivity concerns. The federal clearance process was initiated in June 1991 with the submission of draft materials to the National Center for Education Statistics. The final package containing the final set of cognitive items assembled into blocks and questionnaires was submitted in August 1991. Throughout the clearance process, revisions were made in accordance with changes required by the government. Upon approval, the blocks (assembled into booklets) and questionnaires were ready for printing in preparation for the assessment.

2.4 DEVELOPING THE MATHEMATICS ASSESSMENT

2.4.1 Overview

The 1992 NAEP mathematics assessment was used to measure short-term trend from 1990, so the framework used to develop the 1992 assessment was the same as that used in 1990.

As is the case with all other NAEP instruments, the development of both the objectives and the assessment itself involved a broad-based consensus process. Practitioners, policy-makers, and members of the academic community, representative of many diverse constituencies and points of view, designed objectives for the mathematics assessment, proposing goals they believed students should achieve in the course of their education. After careful reviews of the objectives, assessment items were developed that were appropriate to those objectives. All items underwent extensive reviews by specialists in mathematics, measurement, and bias/sensitivity, as well as by members of the Mathematics Instrument Development Committee. Additionally, tasks in the fourth- and eighth-grade assessments that were to be included in the Trial State Assessment were reviewed by state representatives. All new exercises included in the operational assessment were field tested in a systematic fashion.

2.4.2 Development of the Assessment Framework

The mathematics objectives framework was developed under the auspices of the Council of Chief State School Officers through a special NAEP planning project sponsored by the National Science Foundation and the National Center for Education Statistics (NAEP, 1988). The Council of Chief State School Officers established the National Assessment Planning Project to consider issues related to the Trial State Assessment. They also formed the Mathematics Objectives Committee, whose members (nominated by 18 national organizations) included policymakers, teachers, mathematicians, parents, and citizens nominated by 18 national organizations. The committee was charged with the development of objectives for the 1990 mathematics assessment at grades 4, 8, and 12. These objectives also governed the 1992 assessment.

Although those involved in the legislatively mandated consensus development process drew upon the available draft of the *Curriculum and Evaluation Standards for School Mathematics*, developed by the National Council of Teachers of Mathematics (1989), the project involved widespread participation and review. Input came from a variety of sources, including an objectives committee of mathematics educators; a steering committee with 18 members representing policymakers, practitioners, and citizens at large; the mathematics supervisors in the education agencies of all 50 states for review by state committees; and mathematics scholars and National Center for Education Statistics staff.

More specifically, the development of the framework and objectives proceeded in the following fashion: The Mathematics Objectives Committee developed a draft framework, set of objectives, and set of sample items, which were distributed to the mathematics supervisor in each of the 50 State Education Agencies. These supervisors convened a panel that reviewed the draft and returned comments and suggestions to the project staff. Copies of the draft were also sent to 25 mathematics educators and scholars for review. The Mathematics Objectives Committee incorporated the recommendations made and formulated their final recommendations, which were approved by the National Assessment Planning Project Steering Committee at the Council of Chief State School Officers.

The framework and objectives were then submitted to the National Center for Education Statistics, which forwarded them for review to the Assessment Policy Committee, a panel that advised on NAEP policy at that time. The Assessment Policy Committee approved the objectives with minor provisions about the feasibility of full implementation.[2] The framework and objectives were refined by NAEP's Mathematics Instrument Development Committee, reviewed by the Task Force on State Comparisons, and resubmitted to the National Center for Education Statistics for adoption. Later, the framework was endorsed by the National Assessment Governing Board for use in both the 1990 and 1992 mathematics assessments.

2.4.3 Framework and Assessment Design Principles

Two principles emerged during the discussions of the Mathematics Objectives Committee and became the basis for structuring the framework and objectives for the assessment. The first principle was that a national assessment, designed to provide state-level comparisons, should not measure only those topics and skills already included in the objectives of all states nor be geared to the *least common denominator* of student preparation. The second principle was that the assessment should not be used to steer instruction toward one particular pedagogical or philosophical viewpoint to the exclusion of others that are widely held.

The objectives development was also guided by several other considerations: that the assessment should 1) reflect many of the states' curricular emphases and objectives; 2) reflect what various scholars, practitioners, and interested citizens believe should be included in the curriculum; and 3) maintain some of the content of prior assessments to permit reporting of trends in performance. Accordingly, the committee gave attention to several frames of reference:

- states' goals and concerns, as reflected through analyses of state mathematics curriculum guides and the recommendations of state mathematics specialists;

- a report on "Issues in the Field," based on telephone interviews with leading mathematics educators, and a draft assessment framework provided by a subcommittee of the Mathematics Objectives Committee;

- the draft of the *Curriculum and Evaluation Standards for School Mathematics*, developed by the National Council of Teachers of Mathematics through intensive work by leading mathematics educators in the United States (NCTM, 1989); and

- the design of the 1986 mathematics assessment (NAEP, 1987). The framework for the 1986 NAEP mathematics assessment had 35 cells—seven content and five process areas. Because there were so many cells, the weightings assigned to some of the cells in the 1986 framework did not result in a sufficient number of items to provide reliable measures of students' knowledge and skills. As a result, it was decided that the outline or matrix guiding the development of the 1990 mathematics assessment had to be simplified—rather than having a large number of cells, necessary complexity could be reflected through the designation of specific abilities and topics in each content area.

---

[2] This action was contained in a statement issued by the Assessment Policy Committee's Executive Committee on April 29, 1988. The recommendations were ratified by the full committee on June 18, 1988, with two stipulations: that the objectives be so weighted as to permit reporting on trends in performance; and, with regard to the use of calculator-active items and open-response questions, that the assessment be developed within the resources available for its administration.

2.4.4 Framework for the 1992 Assessment

The framework adopted for the 1990 mathematics assessment (and therefore also for the 1992 mathematics assessment) is organized according to three mathematical abilities and five content areas. The mathematical abilities assessed were conceptual understanding, procedural knowledge, and problem solving. Content was drawn primarily from elementary and secondary school mathematics up to, but not including, calculus. The content areas assessed were numbers and operations; measurement; geometry; data analysis, statistics, and probability; and algebra and functions. Figures 2-3 and 2-4 describe, respectively, the three mathematical abilities and the five mathematical content areas.

The assignment of the percentages of assessment items to be devoted to each mathematical ability and content area was an important component of the framework development because such weighting reflects the importance or value given to each area at each grade level. The National Assessment Planning Project wanted to create an assessment that would be forward-thinking and could lead instruction; thus, they decided to give more emphasis than in previous assessments to problem solving, geometry, and algebra and functions, and less to numbers and operations.

The distribution of items by mathematical ability and mathematical content area for each grade as defined in the framework is provided in Table 2-6 and Table 2-7.

2.4.5 Developing the Cognitive Items

The 1992 mathematics assessment was designed to estimate trends from 1990 in national performance at all three grade levels and at grade 8 for states and jurisdictions that participated in both the 1990 and 1992 Trial State Assessments. However, while the goal of measuring change over time was the primary goal of the assessment developers, the cognitive questions used in the instrument evolved in several significant ways.

The 1992 assessment included a broad range of questions that required students to solve problems in both constructed-response and multiple-choice formats, to provide responses using protractors/rulers, to use calculators (four-function at grade 4 and scientific at grades 8 and 12), and to employ manipulative geometric shapes for computational purposes. The 1992 assessment also included extended constructed response questions that allowed students to demonstrate—in writing and diagrams—their mathematical and problem-solving abilities. Compared to the 1990 assessment, greater emphasis was placed on the ability of students to construct responses, and the percentage of multiple-choice questions was reduced. The special component of the assessment in which students are led by audiotape through a series of tasks designed to measure their estimation skills was continued from 1990. Special study blocks focused on specific mathematical areas at grades 8 and 12 were also included in the 1992 assessment.

Figure 2-3

Description of Mathematical Abilities

The following three categories of mathematical abilities are not to be construed as hierarchical. For example, problem solving involves interactions between conceptual knowledge and procedural skills, but what is considered complex problem solving at one grade level may be considered conceptual understanding or procedural knowledge at another.

Conceptual Understanding

Students demonstrate conceptual understanding in mathematics when they provide evidence that they can recognize, label, and generate examples and counterexamples of concepts; can use and interrelate models, diagrams, and varied representations of concepts; can identify and apply principles; know and can apply facts and definitions; can compare, contrast, and integrate related concepts and principles; can recognize, interpret, and apply the signs, symbols, and terms used to represent concepts; and can interpret the assumptions and relations involving concepts in mathematical settings. Such understandings are essential to performing procedures in a meaningful way and applying them in problem-solving situations.

Procedural Knowledge

Students demonstrate procedural knowledge in mathematics when they provide evidence of their ability to select and apply appropriate procedures correctly, verify and justify the correctness of a procedure using concrete models for symbolic methods, and extend or modify procedures to deal with factors inherent in problem settings. Procedural knowledge includes the various numerical algorithms in mathematics that have been created as tools to meet specific needs in an efficient manner. It also encompasses the abilities to read and produce graphs and tables, execute geometric constructions, and perform noncomputational skills such as rounding and ordering.

Problem Solving

In problem solving, students are required to use their reasoning and analytic abilities when they encounter new situations. Problem solving includes the ability to recognize and formulate problems; determine the sufficiency and consistency of data; use strategies, data models, and relevant mathematics; generate, extend, and modify procedures; use reasoning (i.e., spatial, inductive, deductive, statistical, and proportional); and judge the reasonableness and correctness of solutions.

Figure 2-4

Description of Mathematical Content Areas

Numbers and Operations

This content area focuses on students' understanding of numbers (whole numbers, fractions, decimals, and integers) and their application to real-world situations, as well as computational and estimation situations. Understanding numerical relationships as expressed in ratios, proportions, and percents is emphasized. Students' skills in estimation, mental computation, use of calculators, generalization of numerical patterns, and verification of results are also included.

Measurement

This content area focuses on students' ability to describe real-world objects using numbers. Students are asked to identify attributes, select appropriate units, apply measurement concepts, and communicate measurement-related ideas to others. Questions are included that require an ability to read instruments using metric, customary, or nonstandard units with emphasis on precision and accuracy. Questions requiring estimation; measurements; and applications of measurements of length, time, money, temperature, mass/weight, area, volume capacity, and angles are also included under this content area.

Geometry

This content area focuses on students' knowledge of geometric figures and relationships and on their skills in working with this knowledge. These skills are important at all levels of schooling as well as in practical applications. Students need to be able to model and visualize geometric figures in one, two, and three dimensions and to communicate geometric ideas. In addition, students should be able to use informal reasoning to establish geometric relationships.

Data Analysis, Statistics, and Probability

This content area focuses on data representation and analysis across all disciplines and reflects the importance and prevalence of these activities in our society. Statistical knowledge and the ability to interpret data are necessary skills in the contemporary world. Questions emphasize appropriate methods for gathering data, the visual exploration of data, and the development and evaluation of arguments based on data analysis.

Algebra and Functions

This content area is broad in scope, covering a significant portion of the grade 9-12 curriculum, including algebra, elementary functions (pre-calculus), trigonometry, and some topics in discrete mathematics. For the fourth grade, and in part at grade 8, algebraic and functional concepts are treated in more informal, exploratory ways. Proficiency in this content area requires both manipulative facility and conceptual understanding; it involves the ability to use algebra as a means of representation and to use algebraic skills and concepts as problem-solving tools. Functions are viewed not only terms of algebraic formulas, but also in terms of verbal descriptions, tables of values, and graphs.

Table 2-6

NAEP Mathematics Framework:
Percentage Distribution of Items by Grade and Ability

Mathematical Ability	**Grade**		
	4	**8**	**12**
Conceptual Understanding	40%	40%	40%
Procedural Knowledge	30%	30%	30%
Problem Solving	30%	30%	30%

Table 2-7

NAEP Mathematics Framework:
Percentage Distribution of Items by Grade and Content Area

Mathematical Content Area	**Grade**		
	4	**8**	**12**
Numbers and Operations	45%	30%	25%
Measurement	20%	15%	15%
Geometry	15%	20%	20%
Data Analysis, Statistics, and Probability	10%	15%	15%
Algebra and Functions	10%	20%	25%

All of the constructed-response items were designed to provide an extended view of students' mathematical knowledge and skills. Building on recommendations from the report of the Council of Chief State School Officers, the NAEP Mathematics Instrument Development Committee suggested that constructed-response items be used to assess objectives in the framework that are best measured using such types of items (e.g., the ability to articulate mathematical ideas, draw figures, or generalize function relationships). About half of the constructed-response questions required short answers; the other half, including the extended constructed-response questions, required the ability to formulate and demonstrate more detailed problem-solving skills.

To permit linking to the 1990 assessment, some of the items used in 1990 were used again in 1992. At grade 4, 57 items that were used in the 1990 program were carried forward to the 1992 program (16 short constructed-response items and 41 multiple-choice items). At grade 8, 76 items were used again (23 short constructed-response items and 53 multiple-choice items) and at grade 12, 80 items were reused (24 short constructed-response and 56 multiple-choice items). In addition, the questions in the paced tape estimation blocks had been previously used in the 1990 assessment. The rest of the items used in the 1992 program were newly created. In total, the 1992 assessment, which was nearly doubled in scope, included many more items than did the 1990 assessment.

A carefully developed and proven series of steps were used to create all the questions in the assessment. These steps are described earlier in this chapter under section 2.2.

The development of extended constructed-response questions followed slightly different and more stringent procedures than were followed for all other items. A committee of mathematics educators from elementary and secondary schools, colleges, and state education agencies met early in 1991, and worked with ETS/NAEP mathematics test development staff to develop extended constructed-response items and scoring guides. These items were carefully reviewed according to the procedures required by the ETS *Standards for Quality and Fairness* (ETS, 1987), including content and sensitivity reviews.

Twelve items at each grade level were field tested in May 1991 in urban, suburban, and rural school districts in New Jersey and Pennsylvania. Each student was administered two extended constructed-response items and each item was given to approximately 50 to 100 students. ETS/NAEP mathematics test development staff scored the extended constructed-response items at a special two-day scoring session. Based on the distribution of scores and on the content specifications, the final set of extended constructed-response items was selected by ETS/NAEP mathematics test development staff and reviewed by the Mathematics Instrument Development Committee. These items were included as the last item in the appropriate blocks.

2.4.6 Distribution of Assessment Items

The 1992 mathematics assessment included 178 cognitive items at grade 4, 205 cognitive items at grade 8, and 201 cognitive items at grade 12. This information, along with data on overlap between grades, is summarized in Table 2-8. Please note that the numbers in this table refer to the number of questions administered, as opposed to those included in the analysis; several items were excluded from the analysis after administration.

Table 2-8
Total Number of Cognitive Questions in the 1992 Mathematics Assessment

Use of Questions	Number of Questions		
	Grade 4	Grade 8	Grade 12
Grade 4 only	88		
Grades 4 and 8 only	55		
Grade 8 only		58	
Grades 8 and 12 only		57	
Grade 12 only			109
Grades 4, 8, and 12	35		
Total per grade	178	205	201
Number short constructed-response	54	59	58
Number extended constructed-response	5	6	6
Number regular multiple-choice	99	118	115
Number estimation multiple-choice	20	22	22

Table 2-9
Distribution of Mathematics Questions Across Content and Ability Areas

Content Area	Grade 4		Grade 8		Grade 12	
	Target	Actual	Target	Actual	Target	Actual
Numbers and Operations	45%	40%	30%	32%	25%	24%
Measurement	20%	20%	15%	17%	15%	16%
Geometry and Spatial Sense	15%	17%	20%	20%	20%	18%
Data Analysis, Statistics, and Probability	10%	12%	15%	15%	15%	16%
Algebra and Functions	10%	11%	20%	16%	25%	26%
Ability Area						
Conceptual Understanding	40%	40%	40%	37%	40%	39%
Procedural Knowledge	30%	20%	30%	24%	30%	29%
Problem Solving	30%	40%	30%	39%	30%	32%

In addition to the questions described in Table 2-8, special study blocks on word problems, pre-algebra, and data analysis were included in the national assessment. These blocks included 17 questions at grade 4, 17 questions at grade 8, and 24 questions at grade 12.

In crafting the assessment, all efforts were made to meet targets specified in the assessment framework. Table 2-9 shows the percentage of items in the content and ability areas in the 1992 assessment.

2.4.7 Mathematics Background Questionnaires

As was mentioned above, NAEP includes both general background questionnaires administered to participants in all subject assessments and subject-specific questionnaires for both students and teachers. The development and nature of the general questionnaires is discussed later in this chapter. However, it is worth noting here that members of the Mathematics Instrument Development Committee were consulted on the appropriateness of the issues addressed in all questionnaires that might relate to mathematics instruction and achievement. All background questions were also submitted to extensive field-testing.

In addition to the three blocks of cognitive questions, each participating student answered two five-minute sets of background questions, one general and one on instructional, behavioral, and attitudinal variables related to mathematics. A one-minute questionnaire was given to students at the end of each booklet to determine students' motivation in completing the assessment and their familiarity with assessment tasks.

Mathematics Student Questionnaires

Three categories of information were represented in the five-minute **student mathematics questionnaire** (18 questions at grade 4 and 23 questions at grades 8 and 12): time spent on task and mathematics coursework, the nature of students' mathematics instruction, and students' enjoyment of and confidence in their abilities in mathematics and their perceptions of the usefulness of the discipline to their present and future lives. This questionnaire was the second section in every student assessment booklet.

Mathematics Teacher Questionnaires

To supplement the information on instruction reported by students, the mathematics teachers of the fourth- and eighth-grade students participating in the assessment were asked to complete a mathematics teacher questionnaire about their instructional practices, teaching backgrounds, and characteristics. The teacher questionnaires contained two parts. The first part pertained to the teachers' background and general training. The second part pertained to specific training in teaching reading and the procedures the teacher used for *each class* containing an assessed student.

The Teacher Questionnaire, Part I: Background and Training (23 questions at grade 4 and 32 questions at grade 8) included questions pertaining to gender, race/ethnicity, years of teaching experience, certification, degrees, major and minor fields study, coursework in education, coursework in subject area, in-service training, extent of control over classroom, instruction, and curriculum, and availability of resources for their classroom.

The **Teacher Questionnaire, Part II: Class by Class Mathematics Information** (40 questions at grade 4 and 42 questions at grade 8) pertained to the procedures the teacher used for *each class* containing an assessed student and included questions on the ability level of students in the class, whether students were assigned to the class by ability level, time on task, homework assignments, frequency of instructional activities used in class, instructional emphasis given to the topics and skills covered in the assessment, and use of particular resources.

2.4.8 Development of Final Forms

The field tests for the multiple-choice and short constructed-response items were conducted in February 1991 in 22 states, the District of Columbia, and the Virgin Islands. The intent of the field test was to try out the items and procedures and to give the states and the contractors practice and experience with the proposed materials and procedures. As was mentioned earlier, the extended constructed-response questions were developed and field tested on a different schedule. At grade 4, 172 new exercises were field tested; 101 of these questions were used, along with 77 items from the 1990 assessment, in the operational assessment. At grade 8, 150 new items were field tested; 107 of these were used, along with 77 trend questions, in the operational assessment. At grade 12, 150 exercises were field tested; 99 of these were used in conjunction with 102 trend questions in the operational assessment. About 500 to 600 responses were obtained for each mathematics item in the field test.

The field test data were scored and analyzed in preparation for meetings with the Mathematics Instrument Development Committee and the Background Questionnaire Committee. Using item analysis procedures, which provide a variety of statistics about each item in the field test (including p-values, biserial correlations, and item characteristic curves), committee members, ETS test development staff, and NAEP/ETS staff reviewed the materials to determine

- the most appropriate items for use in the 1992 assessment in accordance with content specifications (that they met the content and ability specifications in the framework) and statistical attributes (that their biserial correlation was not less than 0.20);

- the need for revisions to items that lacked clarity or had ineffective item statistics; and

- appropriate timing for assessment items.

Once the pool of newly created items was established, the items were assembled into nine different "blocks" (15-minute sections established according to statistical guidelines developed at the beginning of the process).[3] The new blocks were assembled taking into account the speededness data from the field test and the fact that extended constructed-response items would be included in certain of the blocks.

Once the total set of items had been selected and assembled into blocks, all items and blocks were reviewed again by ETS/NAEP staff for content, measurement, and sensitivity concerns. In addition, another meeting of representatives from State Education Agencies was convened to review the field test results and final set of items. The federal clearance process was initiated in August 1991 with the submission of materials to the National Center for Education Statistics. Revisions were made in accordance with changes required by the National Center for Education Statistics and the National Assessment Governing Board and the final clearance package was approved in September 1991.

2.5 DEVELOPING THE WRITING ASSESSMENT

2.5.1 Overview

The objectives for the 1992 NAEP writing assessment were developed through a consensus project managed by the National Assessment Governing Board. Educators, policymakers, and scholars participated in developing the framework that governed the development of the assessment. In addition to extensive review by ETS content, sensitivity, and editorial specialists, the writing assessment prompts were reviewed by a wide variety of external experts, including the writing educators and scholars on the Instrument Development Committee. All writing prompts in the operational assessment were systematically field tested.

2.5.2 Development of the Assessment Framework

The 1992 *Writing Framework* was developed through a national consensus process managed by the National Assessment Governing Board. In this project, a number of steps were undertaken:

- The National Assessment Governing Board appointed a Writing Consensus Committee (Appendix A) that was charged with overseeing the development of the writing framework. The panel was composed of writers, teachers, policymakers, and representatives of professional organizations and the corporate community.

- The National Assessment Governing Board sent letters asking for opinions on the development of a writing framework to educators, policymakers, political officials, and business representatives around the country. Responses were used by the Writing Consensus Committee to help in the framework development.

---

[3] In total, there were 13 blocks at each grade level, nine newly created blocks and four trend blocks that had been used in the 1990 mathematics assessment.

- Drafts of the framework were reviewed by large groups of people, representatives of professional associations, and members of the National Assessment Governing Board.

During this development process both the National Assessment Governing Board and the Writing Consensus Committee sought advice from the widest possible range of interested publics, and reviewed innovative writing assessment work being conducted around the country.

2.5.3 Framework and Assessment Design Principles

The writing framework is based on the notion that all students should learn to be effective writers, and that it is important to measure both their writing competency and the attitudinal, instructional, and background variables that may affect their writing achievement. Six core design objectives (NAGB, 1992, p. 9) governed the development of the writing framework and the assessment:

- "Students should write for a variety of purposes: informative, persuasive, and narrative." Different types of writing involve different strategies and conventions, and it is important in NAEP to measure abilities in a range of major writing purposes.
- "Students should write on a variety of tasks, and for many different audiences." Different writing tasks and different audiences put varied demands on writers, so NAEP should specify a wide variety of writing tasks and diverse audiences in writing exercises.
- "Students should write from a variety of stimulus materials, and within various time constraints." Because stimuli and specific assignments may affect writing, it is essential that NAEP use a wide variety of stimulus materials and writing prompts in the assessment, so that measurement of writing skill is not skewed by a particular type of stimulus.
- "Students should generate, draft, revise, and edit ideas and forms of expression in their writing." Writing is, at its core, an iterative process. It is therefore important to allow enough time in individual exercises for students to plan, compose, and edit written responses.
- "Students should display effective choices in the organization of their writing. They should include detail to illustrate and elaborate their ideas, and use appropriate conventions of written English." To effectively measure writing, it is important that NAEP exercises require students to organize arguments, and to include incorporate supporting material in a given piece of writing.
- "Students should value writing as a communicative activity." In the context of NAEP, it is imperative that the assessment include background variables regarding attitudes students hold toward writing. It is also important that in NAEP analyses these variables be related to writing achievement.

2.5.4 Framework for the 1992 Assessment

The *NAEP Writing Framework* identifies three primary purposes for writing—informative, narrative, and persuasive. Informative writing "focuses primarily on the subject matter element in communication" (NAGB, 1992, p. 9) and is used to share knowledge and to convey ideas. Narrative writing encourages students to incorporate their imagination and creativity into the production of stories or personal essays. Persuasive writing focuses on the reader with the primary aim of influencing "others to take some action or bring about change" (NAGB, 1992). These three writing purposes were addressed at all three grades assessed, with students at grades 4, 8, and 12 responding to two 25-minute tasks or, at grades 8 and 12, one 50-minute writing task. Some tasks were given at more than one grade.

To evaluate the kinds of ways students plan what they will write, NAEP provided a space for students to engage in pre-writing activities. Their pre-writing methods were recorded in seven categories, ranging from diagrams or outlines to complete first drafts. However, since the assessment context provides little opportunity to review and revise one's work, students' responses to assessment tasks were viewed as first-draft writing and evaluated accordingly.

The writing framework also explicitly discusses the ways in which the assessment tasks should be scored. Students' responses to each writing task were evaluated by trained raters who used primary-trait analysis. The scoring guidelines defined six successive levels of task accomplishment: response to topic; undeveloped response to task; minimally developed response; developed response; elaborated response; and extensively elaborated response. Each task had a unique scoring guide. However, all the scoring guides were based on the primary-trait framework presented in Figure 2-5.

The 6-point scoring rubric represents a departure from the 1988 assessment, where a 4-point scale was used. It was judged that moving to a 6-point scale would provide more data about how American students write.

2.5.5 Developing the Cognitive Items

As was the case with all NAEP assessments administered in 1992, the writing assessment development effort was managed by Educational Testing Service, whose assessment development specialists worked closely with the 1992 Writing Task Development Committee. The Committee, composed of experts in writing education and assessment, played a central role in developing, reviewing, and approving all exercises and stimulus materials.

The cognitive portion of the writing assessment included only constructed-response exercises. Questions and tasks were designed to measure students' abilities to write for a variety of purposes and to a diverse set of audiences. To accomplish these goals, a wide variety of stimulus materials were used in the assessment. The first step in the development effort was the identification of appropriate stimulus materials that would allow the construction of tasks that would, in aggregate, measure the range of writing outcomes described in the framework.

A carefully developed and proven series of steps was used to create the assessment items. These steps are described earlier in this chapter in section 2.2.

Figure 2-5

NAEP Primary Trait Scoring Guidelines

Primary trait: [In each scoring guide, a specific trait is defined here, corresponding to the requirements of the given task. For example, the primary trait for a narrative task would be *quality of narrative*; for an informative task, *quality of description (clarity and use of detail)*; and for a persuasive task, *quality of argument (clarity of perspective and level of support)*].

Scoring rationale: [In each scoring guide, a rationale is provided here, summarizing the task and explaining the basis for scoring papers. The scores ranged from 6 to 0, with 0 being the lowest valid score.]

6 **Extensively elaborated**. In these papers, students write a well developed, detailed, and well written response to the task. They show a high degree of control over the various elements of writing. These responses may be similar to "5" responses, but they are better organized, more clearly written, and less flawed.

5 **Elaborated**. In these papers, students write a well developed and detailed response to the task. They may go beyond the requirements of the task.

4 **Developed**. In these papers, students provide a response to the task that contains necessary elements. However, these papers may be unevenly developed.

3 **Minimally developed**. In these papers, students provide a response to the task that is brief, vague, or somewhat confusing.

2 **Undeveloped response to task**. In these papers, students begin to respond to the task, but they do so in a very abbreviated, confusing, or disjointed manner.

1 **Response to topic**. In these papers, students respond to some aspect of the topic but do not appear to have fully understood the task. Or, they recopy text from the prompt.

0 **Not rated.** These papers were blank.

The assessment consisted of different 25-minute and 50-minute "blocks," each consisting of a single writing prompt. Students were asked to respond to either two 25-minute tasks or one 50-minute task within one booklet (50-minute tasks were used at grades 8 and 12 only). At all grades, the assessment included nine 25-minute blocks. In addition, the grade 8 assessment included two 50-minute blocks and the grade 12 assessment three 50-minute blocks. Therefore, the total number of assessment blocks was 9 at grade 4, 11 at grade 8, and 12 at grade 12. The decision to give students either 25 or 50 minutes to respond to a topic was made based on the results of field testing: field test results were analyzed to identify those tasks that provided students with ample information and motivation to justify 50 minutes of assessment time.

A subset of fourth- and eighth-grade students who were assessed in writing were selected to participate in "The Nation's Writing Portfolio," a special project designed to provide information about the type of writing students do in school. The primary writing teachers of sampled students were asked to assist students in selecting three pieces of writing that represented their best writing efforts, a range of writing tasks, and the use of writing process strategies. In addition, students were asked to write a letter to NAEP explaining why they chose their three pieces and what they like about each piece.

2.5.6 Writing Background Questionnaires

The writing framework placed a strong emphasis on measuring the relationship between additional and educational background variables and writing achievement. Therefore, student questionnaires at all grades and a teacher questionnaire at grade 8 were included as important components of the 1992 writing assessment.

NAEP includes both general background questionnaires given to participants in all subjects and subject-specific questionnaires for both students and their teachers. The development of the general background is discussed below. Like the cognitive items, all background questions were submitted to extensive review and field testing.

Students participating in the writing assessment answered a set of general background questions and a set of writing-specific questions. Together, these sets of questions were designed to gather contextual information about students, the nature of their writing instruction and writing practices, and their attitudes toward writing. A one-minute questionnaire was given to students at the end of each booklet to determine their motivation in completing the assessment and their familiarity with assessment tasks. In order to ensure that all grade 4 students understood the questions and had every opportunity to answer them, the three sets of questionnaires were read aloud by administrators as students read along and responded in their booklets.

A teacher questionnaire was administered to the English or language arts teachers of eighth graders participating in the assessment. This questionnaire contained two parts. The first requested information about the teacher, such as race/ethnicity and gender, as well as academic degrees held, teaching certification, training in writing, and ability to get instructional resources. In the second part, teachers were asked to provide information on each class they taught that included one or more students who participated in the assessment. The information included, among other things, the amount of time spent on writing instruction and homework, the extent to which various writing assignments were given, the instructional and grading emphases placed on different aspects of writing, and the use of various instructional approaches (e.g., peer response and computers).

2.5.7 Development of Final Forms

Writing field tests were conducted in February 1991. The intent of the field test was to try out the items and procedures and to give the contractors practice and experience with the proposed materials and procedures. At grade 4, 12 blocks were field tested; at grade 8, 16 blocks were field tested; and at grade 12, 18 blocks were field tested. This yielded 9 blocks at grade 4, 11 at grade 8, and 12 at grade 12. Twenty-two unique writing prompts were used in the operational assessment (some prompts were used for more than one grade). The design required that about 500 responses be obtained to each item in the field test. Scoring of the writing tasks, which was a central component of the writing field test, was conducted at the National Computer Systems Iowa City headquarters during March 1991.

The field-test data were collected, scored, and analyzed in preparation for meetings with the Writing Instrument Development Committee. Committee members, ETS test development staff, and NAEP/ETS staff reviewed the materials. Several objectives guided these reviews: to determine the need for revisions of writing tasks that lacked clarity; to prioritize prompts to be included in the assessment; and to double-check the timing for the writing prompts.

Once the committee had selected the writing prompts, they were rechecked for content, measurement, and sensitivity concerns. The federal clearance process was initiated in June 1991 with the submission of draft materials to the National Center for Education Statistics. The final package containing the final set of cognitive items assembled into blocks and questionnaires was submitted in August 1991. Throughout the clearance process, revisions were made in accordance with changes required by the government. Upon approval, the blocks (assembled into booklets) and questionnaires were ready for printing in preparation for the assessment.

2.6 BACKGROUND QUESTIONNAIRES

As part of NAEP, a series of questionnaires was administered to students, teachers, and school administrators. Similar to the development of the cognitive items, the development of the policy framework and questionnaire items was a consensual process that involved staff work, field testing, and review by external advisory groups.

A Background Questionnaire Committee was appointed to oversee the development of a policy framework and individual questions. The policy framework this panel developed focused on five educational areas: instructional content, instructional practices and experiences, teacher characteristics, school conditions and contexts, and condition outside of school (NAEP, 1992). The panel was also interested in capitalizing on the unique qualities of NAEP and not duplicating other surveys.

The panel recommended that study be focused on areas that addressed the relationship between student achievement and instructional practices. The policy issues, items, and field test results were reviewed by a group of external consultants who identified specific items to be included in the final questionnaires. In addition, members of the Instrument Development Committee in all three subjects reviewed the questionnaires. The questionnaires underwent ETS reviews designed to ensure quality and fairness. All questions were systematically field tested.

2.6.1 Student Questionnaires

In addition to the cognitive questions, the 1992 assessment included two five-minute sets of general and subject-specific background questions designed to gather contextual information about students, their experiences in mathematics, and their perceptions of the subject, and a one-minute set of background questions about the students' motivation regarding the assessment. In many cases the questions used were continued from prior assessments in order to measure change over time.

The **student demographics (common core) questionnaire** included questions about race/ethnicity, language spoken in the home, mother's and father's level of education, reading materials in the home, television watching, homework, and which parents live at home. This questionnaire was the first section in every booklet. The exact content subject specific questionnaires have been discussed

The **student subject-specific questionnaires** covered issues such as time spent studying a specific subject, coursework, the nature of students' instruction, and students' enjoyment of and confidence in their abilities in subject areas and their perceptions of the usefulness of the disciplines to their present and future lives. This questionnaire was the second section in every booklet. This questionnaire was the fourth section in reading and writing booklets, and the second section of all mathematics booklets.

The **motivation questionnaire** asked the students questions about their perceptions of the difficulty of the assessment, and of how well they did on the assessment, and their motivation to do well on the assessment. This questionnaire was the last section in every booklet.

2.6.2 Teacher, School, and Excluded Student Questionnaires

To supplement the information on instruction reported by students, the mathematics teachers of the fourth- and eighth-grade students participating in the mathematics assessment, the reading teachers of fourth-grade students participating in the reading assessment, and the writing instructors of eighth-grade students in the writing assessment were asked to complete a teacher questionnaire about their instructional practices, teaching backgrounds, and characteristics. The teacher questionnaires contained two parts. The first part pertained to the teachers' background and general training. The second part pertained to specific training in teaching reading and the procedures the teacher used for *each class* containing an assessed student.

Because the sampling for the teacher questionnaires was based on participating students, the responses to a particular teacher questionnaire do not necessarily represent all teachers of that subject area at that grade level in the nation. Rather, they represent teachers of the representative sample of students assessed. It is important to note that in this report, as in all NAEP reports, the student is always the unit of analysis, even when information from the teacher or school questionnaire is being reported. Using the student as the unit of analysis makes it possible to describe the instruction received by representative samples of students. Although this approach may provide a different perspective from other studies simply reporting information about teachers or schools, it is consistent with NAEP's goal of providing information about the educational context and performance of students.

The Teacher Questionnaire, Part I: Background and Training included questions pertaining to gender, race/ethnicity, years of teaching experience, certification, degrees, major and minor fields study, coursework in education, coursework in subject area, in-service training, extent of control over classroom, instruction, and curriculum, and availability of resources for their classroom.

The **Teacher Questionnaire, Part II: Class by Class Subject-Specific Information** pertained to the procedures the teacher uses for *each class* containing an assessed student and included questions on the ability level of students in the class, whether students were assigned to the class by ability level, time on task, homework assignments, frequency of instructional activities used in class, instructional emphasis given to the topics and skills covered in the assessment, and use of particular resources.

School Characteristics and Policies Questionnaires were given to the principal or other administrator of each school that participated in NAEP. This questionnaire included questions about background and characteristics of school principals, length of school day and year, school enrollment, absenteeism, drop-out rates, policies about tracking, curriculum, testing practices and use, special priorities and school-wide programs, availability of resources, special services, community services, policies for parental involvement, and school-wide problems.

The **Excluded Student Questionnaire** was completed by the teachers of those students who were selected to participate in the assessment sample but who were determined by the school to be ineligible to be assessed because they either had an Individualized Education Plan (IEP) and were not mainstreamed at least 50 percent of the time, or were categorized as Limited English Proficient (LEP). This questionnaire asked about the nature of the student's exclusion and the special programs in which the student participated.

Schools were permitted to exclude certain students from the assessment. The same exclusion criteria and rules used in the national assessment were also applied to the Trial State Assessment. Although the intent was to assess all sampled students, students who were identified by school staff as not capable of participating meaningfully were excluded. The NAEP guidelines for exclusion are intended to assure uniformity of exclusion criteria from school to school as well as from state to state.

Chapter 3

SAMPLE DESIGN

Keith F. Rust and Leslie Wallace

Westat, Inc.

The samples for the 1992 NAEP assessment were selected using a complex multistage sample design involving the sampling of students from selected schools within 94 selected geographic areas, called primary sampling units (PSUs), across the United States.

The sample design had four stages of selection:

1) selection of geographic PSUs (counties or groups of counties);
2) selection of schools within PSUs;
3) assignment of session types to schools; and
4) selection of students for session types within schools.

The samples were drawn for the three different age classes[1], and for each age class the samples were of two distinct types. The first type consisted of the cross-sectional or "main" samples, while the second type consisted of the long-term trend samples. The populations surveyed with each of these sample types are defined in Table 1-1. Separate samples of schools were required for the long-term trend samples and main samples, because of various differences in the calendar period for test administration, the format of the administration, and, in the case of age class 17, the grade and age definition of the population of interest.

In addition to representing the respective populations as a whole, for the main samples there was oversampling of private schools, and of public schools with moderate or high enrollment of Black or Hispanic students (see section 3.2). This oversampling was undertaken to increase the sample sizes of private school students and minority students, so as to increase the reliability of estimates for these groups of students.

---

[1] The term "age class" is used in this report when it is appropriate to discuss one of the three student cohorts in a general way (not necessarily in reference to a specific sample). For the 1992 assessment, age class 1 refers to age 9 and age 9/grade 4 students, age class 2 refers to age 13 and age 13/grade 8 students, and age class 3 includes the age 17, age 17/grade 11, and age 17/grade 12 students.

The overall assessment period fell into three time periods—fall, winter, and spring. Not all assessment components were conducted in each time period. Table 3-1 shows the relationship between the various sample components and the assessment periods. The sizes of the PSU and school samples and the procedures for their selection were determined by the assessment period, as well as by the population to be surveyed and the method of administration in each case.

Table 3-1
Assessment Type by Age Class and Assessment Period

Age Class	**Fall 10/7/91 - 12/13/91**	**Winter 1/6/92 - 3/13/92**	**Spring 3/16/92 - 5/15/92**
9	—	Main assessment Long-term trend assessment	—
13	Long-term trend assessment	Main assessment	—
17	—	Main assessment	Long-term trend assessment

The age class 9 and age class 13 long-term trend samples used the same school and student eligibility requirements as the respective main samples. Nevertheless, special trend samples were required because:

1) The conditions for administration of the assessment varied considerably between the main sample and long-term trend sample sessions.

2) The need in the long-term trend samples for four distinct session types for age class 9 and 13 and three age class 17, together with the need for five distinct session types for the main samples for age class 9 and four for age class 13 and 17, made it infeasible to conduct both main sample sessions and long-term trend sessions in a given school.

3) For age class 13, the main samples were conducted at an inappropriate time of the year for trend purposes, so that a distinct sample of schools was needed to undertake the long-term trend assessments in the fall of 1991.

A separate sample of schools was required for the long-term trend sessions and the main sessions for age class 17 primarily because the definitions for student eligibility, based on age and grade, differed substantially between the two samples, even though the same population of schools was surveyed in each case. Conditions of administration also varied somewhat, and there were up to four distinct trend session types and five distinct main sample session types per age class. Thus, it was not feasible to conduct main sample and long-term trend sample sessions within a single school.

This chapter gives details of the sample selection procedure, and information on the results of the sampling process. Fuller details are given in the report *1992 National Assessment of Educational Progress Sampling and Weighting Procedures, Final Report* (Wallace & Rust, 1994).

3.1 PRIMARY SAMPLING UNITS

In the first stage of sampling, the United States (the 50 states and the District of Columbia) was divided into geographic primary sampling units (PSUs). Each PSU met a minimum size requirement (a population of at least 60,000 in the 1980 Census) and comprised either a metropolitan statistical area (MSA), a single county, or (more usually in the case of nonMSA PSUs) a group of contiguous counties. In the case of New England MSAs, which are not formed from whole counties, the corresponding New England County Metropolitan Areas, which are defined in terms of whole counties, were designated as the PSUs. The New York City MSA was divided along county/borough lines into three PSUs for reasons of administrative and sampling convenience. Each PSU was contained entirely within one of the four regions defined in Table 3-2. These regions were used to stratify the sample of PSUs, ensuring that each region was adequately represented in the various assessment samples.

Table 3-2
Geographic Regions Used for Stratification

Northeast	**Southeast**	**Central**	**West**
Connecticut	Alabama	Illinois	Alaska
Delaware	Arkansas	Indiana	Arizona
District of Columbia	Florida	Iowa	California
Maine	Georgia	Kansas	Colorado
Maryland	Kentucky	Michigan	Hawaii
Massachusetts	Louisiana	Minnesota	Idaho
New Hampshire	Mississippi	Missouri	Montana
New Jersey	North Carolina	Nebraska	Nevada
New York	South Carolina	North Dakota	New Mexico
Pennsylvania	Tennessee	Ohio	Oklahoma
Rhode Island	Virginia*	South Dakota	Oregon
Vermont	West Virginia	Wisconsin	Texas
Virginia*			Utah
			Washington
			Wyoming

---

* That part of Virginia that is part of the Washington, DC-MD-VA metropolitan statistical area is included in the Northeast region; the remainder of the state is included in the Southeast region.

In a few cases an MSA crossed region boundaries. Such MSAs were split into two or more PSUs as necessary (e.g., the Cincinnati OH-KY-IN MSA was split into the Cincinnati

OH-IN PSU in Region 3 and the Cincinnati KY PSU in Region 2). Twelve subuniverses of PSUs were then defined as described below.

The 28 largest PSUs were included in the sample with certainty. An additional six very large PSUs (four from the Southeast and two from the West) that had large proportions of Black students and/or Hispanic students were also included with certainty. The 34 certainty PSUs constituted 32 metropolitan areas, since the New York City MSA was divided into three certainty PSUs. The inclusion of these 34 PSUs in the sample with certainty provided an approximately optimum, cost-efficient sample of schools and students when samples were drawn within them at the required national sampling rate. The representativeness of the sample for minority groups was enhanced by ensuring that these PSUs were included in the sample, since these minority groups are relatively heavily represented within these certainty PSUs. The remaining smaller PSUs were not guaranteed to be selected for the sample. These were grouped into a number of noncertainty strata (so called because the PSUs in these strata were not included in the sample with certainty), and sample PSUs were selected from each stratum.

The PSUs were classified into four regions, each containing about one-fourth of the U.S. population. These regions were defined primarily by state (Table 3-2). In each region, noncertainty PSUs were classified as MSA or nonMSA. In the Southeast and West regions, a PSU was classified as high minority if the combined proportion of the population of Blacks (in the Southeast) and Hispanics (in the West) in the 1980 Census exceeded 20 percent. The resulting major strata, or subuniverses, are shown in Table 3-3.

Table 3-3
The Sampling Subuniverses
and the Number of Noncertainty Strata in Each

	MSA PSUs		**NonMSA PSUs**	
Region	**Regular Strata**	**High-minority Strata**	**Regular Strata**	**High-minority Strata**
Northeast	8	—	2	—
Southeast	4	6	4	6
Central	8	—	6	—
West	4	6	4	2
Total	24	12	16	8

Within each major stratum (subuniverse), further stratification was achieved by ordering the noncertainty PSUs according to several additional socioeconomic characteristics, yielding 60 strata. The number of such strata formed within each subuniverse is shown in Table 3-3. The strata were defined so that the aggregate of the measures of size of the PSUs in a stratum was approximately equal for each stratum, except for strata in the high-minority subuniverses, in which the aggregate was approximately half that of the regular strata. The size measure used was the population from the 1980 Census. The characteristics used to define strata were the percent minority population, the percentage change in total population since 1970, the per capita educational expenditure, the percent of persons employed in manufacturing (MSA subuniverses only) and the percentages of rural and urban dwellers (nonMSA subuniverses only). One PSU was selected with probability proportional to size from each of the 60 noncertainty strata. That is, within each stratum, a PSU's probability of being the sample selection from that stratum was proportional to its population. Thus the high-minority subuniverses were sampled at approximately twice the rate of the other subuniverses, since they were about half as large. This procedure of oversampling from the high-minority subuniverses was used with the aim of reducing somewhat the level of sampling error for estimates relating to the populations of Black and Hispanic students.

The final sample of 94 PSUs was drawn from a population of about 1,000 PSUs. Primarily because of the use of MSAs as PSUs, PSUs varied considerably as to their probability of selection, since they varied greatly in size. The 34 certainty PSUs consisted of the 26 largest MSAs in the country, based on the 1980 population from the Census, plus six other large MSAs from the Southeast and West regions that had Black or Hispanic populations of more than 20 percent. The 36 selected noncertainty MSA PSUs had probabilities of selection ranging from 0.029 to 0.584, while the 24 selected nonMSA PSUs had probabilities ranging from 0.020 to 0.053. The variations in probability depended upon the size of the PSU (1980 population) and whether or not the PSU was in a high minority subuniverse. Parts of 38 states were included in the 94 selected PSUs.

Samples of 94 PSUs each were drawn at one time for the 1986, 1988, 1990, and 1992 assessments. They were drawn so as to provide for the rotation of the PSUs from one assessment to the next, except that certainty PSUs were retained in each assessment year, and some of the larger noncertainty PSUs were retained for two successive assessment years.

For the long-term trend samples, 64 PSUs were subsampled from the 94 main sample PSUs. The long-term trend samples were much smaller than the main samples and used separate field staff. Fewer PSUs were used for the long-term trend samples to avoid having the sample spread too thinly across PSUs.

The 14 largest main sample certainty PSUs were also included with certainty in the long-term trend samples. Ten additional PSUs were selected systematically and with probability proportional to the 1986 population from the 20 remaining main sample certainties. Finally, 40 PSUs were selected from the 60 noncertainty main sample PSUs so that the overall procedure was equivalent to systematic sampling with probabilities proportional to the 1980 population.

3.2 SCHOOLS FOR MAIN SAMPLES; ASSIGNMENT OF SESSIONS TO SCHOOLS

In the second stage of sampling, the public schools (including Bureau of Indian Affairs schools and Department of Defense schools) and private schools (including Catholic schools) within each of the 94 PSUs were listed according to the grade ranges associated with the three age classes. Table 3-4 shows the numbers of schools included in the various frame components. The population of eligible public schools for each age class was the same for long-term trend and main samples in each case, except that the schools were restricted to the selected PSUs.

Additional private schools were identified and included on the frame for the main samples, but not for the long-term trend samples. Any school having one or more of the eligible grades, and located within an appropriate PSU, was included on the frame of schools (the list of schools from which the samples of schools were drawn) for a given sample. For each age class, only a fraction of one percent of age-eligible students was enrolled in ineligible schools. Each school within the 94 PSUs with a grade in the range of 2 to 12 was included in at least one age class—a total of 37,823 schools. An independent sample of schools was selected for each of the age classes. Thus some schools were selected for assessment of two age classes, and a few were selected for all three.

Table 3-4
Grade Definition of School Eligibility for Frame Inclusion
and Frame Sizes, Main and Long-term trend Samples

Age Class	Frame Included Schools with Any Grade in This Range	Number of Public Schools on Frame*		Number of Private Schools on Frame**	
		Long-term Trend	Main	Long-term Trend	Main
9	2 - 5	14,182	18,497	6,725	10,048
13	6 - 9	13,266	17,304	7,298	10,697
17	9 - 12	3,587	4,745	2,268	4,216
Total	2 - 12	20,172	26,379	7,787	11,444

---

* Public, Bureau of Indian Affairs, and Department of Defense Schools

** Catholic and other private schools

The lists of schools were obtained from several sources. Regular public, Bureau of Indian Affairs, Department of Defense, Catholic, and other private schools were obtained from the 1990 list of schools maintained by Quality Education Data, Inc. (QED).

For the main samples, supplementary lists of private schools were obtained from two sources and added to the QED list of private schools. This supplementation was undertaken because previous studies have revealed that the QED list of private schools is somewhat deficient in its coverage of nonCatholic private schools (Westat, Inc., 1984, Chapter 4). Although the percentage of students in schools not covered by QED is small as a percentage of the total student population (believed to be less than 1 percent), we believed that it was important to attempt to reduce this noncoverage for the main samples, since separate estimates were to be produced for private schools, based on samples of private school students obtained using a relatively high rate of sampling compared to that of public school students for the main sample and private school students in the long-term trend samples and in previous years.

The first supplementary private school listing source used was the Private School Survey (PSS) developed for the National Center for Education Statistics' 1988 School and Staffing

Survey. This list was restricted to a sample of counties selected for the survey. Certain of these counties, generally large in population, were also included, independently by chance, in the NAEP sample PSUs. The schools from such counties were added to the NAEP frame after steps were taken to eliminate duplicates with the QED list of private schools. The second source was a series of lists generated clerically from the yellow pages of telephone directories from metropolitan areas included in the NAEP PSU sample. This list was matched against the other private school sources to eliminate duplicates.

The process of private school list supplementation resulted in the addition of 1,708 schools to the frame, 703 obtained from the School and Staffing Survey list and 1,005 derived from telephone directories. The nature of these listing sources meant that little was known about these schools, in particular the grade span. This meant that a large proportion of these schools were in fact out of scope for a given age class school frame. These 1,708 schools were included in the sampling process, and the 163 that were selected in the samples were initially screened via telephone to establish whether or not the school was in scope. Eighty-one schools were dropped from the sample as a result of this process, and a further 43 were subsequently found to be out of scope by the Westat field supervisor who contacted the school regarding participation in NAEP. Thus it appears that only about 24 percent of the supplementary private schools were in scope for the sample for which they were selected.

For each sample, schools were selected (without replacement) across all PSUs with probabilities proportional to assigned measures of size. Equal measures of size were assigned to schools containing estimates of age/grade-eligible students ranging from 20 to 150 (for age class 9) or 20 to 200 (for age class 13 and age class 17). Schools larger than the indicated maximum size were selected with probabilities proportional to the number of age/grade-eligible students. Schools with fewer than 20 estimated age/grade eligibles were assigned somewhat lower measures of size, and thus lower probabilities of selection, since assessment in these schools involved substantially higher per-student administrative costs.

Each public school with minority (Black and/or Hispanic) enrollment in excess of 10 percent of total enrollment was given double the probability of selection of a public school of similar size in the same PSU with minority enrollment of 10 percent or less. Overall probabilities of selection for such high-minority schools were twice those for other schools of the same size from a given PSU in order to enlarge the sample of Black and Hispanic students, thereby enhancing the reliability of estimates for these groups. For a given overall size of sample, this procedure reduces somewhat the reliability of estimates for all students as a whole and for those not Black or Hispanic.

Each private school was given triple the probability of selection of a low-minority public school of similar size. These greater probabilities of selection were used to ensure adequate samples of private school students in order to allow the derivation of reliable estimates for such students.

The total number of schools selected for each age class was determined to be such that the predesignated student sample sizes would be achieved by selecting all eligible students in a selected school, up to a maximum of 150 (for age class 9) or 200 (for age class 13 and age class 17), allowing for losses due to nonparticipation of selected schools and students and the exclusion of students from the assessment.

This design, with the important exceptions described above, had the goal of yielding a sample of students in a given age or grade with approximately uniform probabilities of selection. The efforts to oversample certain subgroups in the population and the practical constraints on the sample size within each school resulted in some substantial violations of this general goal. The distributions of selection probabilities of the selected students, as reflected in their sampling weights, is discussed in Chapter 10.

For all three age classes, a sample of schools was first drawn for the long-term trend assessments (see section 3.3). These schools were then excluded from the frame when the samples of schools were drawn for the main assessments. Adjustments were made to the sampling weights to reflect the appropriate probabilities of selection to yield unbiased estimates for both long-term trend and main samples.

The QED files do not contain schools that opened between 1990 and the assessment dates. Therefore, special procedures were implemented to be sure that the NAEP assessment represented students in new public schools. Small school districts, which generally contained only one eligible school for a given age class, were handled differently from large school districts, which generally contained more than one eligible school for a given age class. In small school districts, the schools selected for a given age class were thought to contain all students in the district that were eligible for the assessment. Districts containing these schools were asked if other schools with the appropriate grades for the assessment existed, and if so, they were automatically included in the assessment. For large school districts, a district-level frame was constructed from the schools on the QED file that were eligible for one of the national assessments. Then districts were sampled systematically with probabilities proportional to a measure of size. In most cases, the measure of size was total district enrollment, but in very small districts a minimum measure of size was used. Each sampled district was asked to update lists of eligible schools according to information on the QED files. Frames of eligible new schools were then constructed separately for each age class, and separate samples of new schools were selected systematically with probability proportional to eligible enrollment. Five new schools were added to the main samples—three at age class 13 and one each at age classes 9 and 17.

In a few PSUs where school refusals were relatively heavy for a particular sample, substitute school selections were made, replacing the refusals (to the extent feasible) with schools from within the same PSU and similar in size, affiliation (public, Catholic, or other private), grade span, and minority composition. The goal of this procedure was to maintain the student sample sizes needed, while keeping variance and nonresponse bias at acceptable levels. Table 3-5 shows the number of in-scope schools selected, cooperating, and substituted, in each of the school samples. The participation rates given are based on the initially selected sample of schools. These response rates are comparable with those of assessments conducted since 1980. Note that since the response rates quoted do not include the substitute selections, the potential for nonresponse bias is likely to be a little less than these rates would indicate. This is because the substitute selections were chosen based on their similarity to the initially refusing selections. Schools selected for NAEP that had also been selected for the National Educational Longitudinal Survey (NELS) were not contacted for NAEP. Since probability sampling methods were used in NELS, these schools are considered a random subset of the NAEP sample, so that

Table 3-5

School Sample Sizes, Refusals, and Substitutes for the Main Samples

Status	Age Class 9	Age Class 13	Age Class 17	Total	Public*	Private**
Selected, in scope	619	735	605	1,959	1,133	826
Refusals	84	108	112	304	156	148
Participation rate of originally selected schools	86%	85%	81%	84%	86%	82%
1990 participation rate	88%	87%	81%	86%	87%	84%
NELS schools	0	5	71	76	62	14
Participating, no eligibles enrolled	12	44	29	85	41	44
Substitutes participating	4	4	4	12	8	4
Final assessed sample	527	587	468	1,582	944	638

---

* Public, Bureau of Indian Affairs, and Department of Defense schools

** Catholic and other private schools

no bias is introduced by not contacting them. Therefore, they were not counted when calculating response rates. Instead, they were accounted for in weighting with a school nonresponse adjustment.

The considerable numbers of schools selected with no eligible students enrolled resulted primarily from the fact that, for example, for age 13/grade 8, some schools with grades 6, 7, or 9, but no grade 8, were sampled. Such schools had a reasonable chance of containing some age 13 students. Often they did have a number of eligible students, but sometimes they had none. Because of the grade structure of schools, this occurred most often for age 13/grade 8. In addition, many supplemental private schools turned out to be ineligible since they were added to the frame without any knowledge of their grade spans.

A school characteristics and policies questionnaire was mailed to every sampled school by Westat before the assessment. The Westat supervisor then collected the questionnaires and returned them to ETS. The school characteristics and policies questionnaire is described in Chapter 4.

A school principal's questionnaire, distributed to each sampled school by Westat before the assessment, was used to refine the estimate of the age/grade-eligible students and to determine in part the size and type of community (STOC) codes (see Rust, Burke, & Fahimi, 1992).

Five different session types were administered at age class 9 and four were administered at each of age classes 13 and 17. The four session types common to all age classes were print-administered reading, print-administered writing, print-administered mathematics, and tape-administered mathematics. A print-administered mathematics session involving calculators was also included at age class 9 to test the effects of using a calculator different from that used in a previous assessment. These session types were assigned among the selected schools found to be in-scope at the time of assigning sessions.

First, the minimum proportion of sampled students within a school who could be assigned to a single session, without that session's being unduly small, was established. Thus, for schools with few eligible students, all students were to be assigned to a single session, to be one of the session types above. In large schools at age classes 13 and 17, where it was anticipated that 200 students would be selected, this proportion was set at one-sixth. For large schools at age class 9, where it was anticipated that 150 students would be selected, this proportion was set at one-fifth. Intermediate proportions were set for schools in intermediate size (see Rust, Burke, & Fahimi, 1992). Session types were then assigned to schools with three aims in mind. The first was to distribute students to the different session types, across the whole sample (and so far as possible, proportionately within each PSU), for each age class, so that the target numbers of assessed students would be achieved. The second was to maximize the number of different session types that were administered within a given selected school, without violating the minimum proportions within a single session (discussed above). The third was to give each student selected for the main sample for an age class an equal chance of being selected for a given session type. Thus, overall and as far as possible within each PSU, 30 percent of students at age class 9 were assigned to print-administered writing, 26.7 percent to print-administered reading, 6.7 percent to the print-administered mathematics calculator bridge and 6.7 percent to tape-administered mathematics.

Some schools failed to participate after having been assigned session types. Often also such schools were not replaced by substitute selections. Thus two types of school-level nonresponse, designated as school and session, were needed to address the consequences of not contacting NELS schools (discussed earlier) and of nonresponse after session assignments were made. Adjustments to the sample weights were made for each of these nonresponse components, as discussed in Chapter 10.

The procedure was intended to ensure that each session type was conducted within each PSU. The relatively small proportion of schools assigned to conduct the tape-administered mathematics and calculator bridge sessions, and the fact that school-level nonresponse occurred after the assignment of sessions, meant that this was not always possible to achieve. The use of this procedure, however, helped to ensure that the different session types were spread among PSUs to the maximum extent feasible in practice.

3.3 SELECTION OF SCHOOLS FOR LONG-TERM TREND SAMPLES; THE ASSIGNMENT OF SESSIONS TO SCHOOLS

Long-term trend sample schools were selected for each age class from the 64 long-term trend PSUs. This was a change from 1990 when long-term trend samples were selected from all 94 PSUs, but similar to earlier years, when long-term trend samples were selected from a subset of PSUs. The different administrative procedures required for the main and long-term trend assessments, coupled with the smaller sample sizes for the long-term trend assessments made it desirable to conduct the long-term trend assessments in a subset of PSUs using field staff independent of that used in the main assessments.

The sample of schools was drawn for the long-term trend samples in a manner very similar to that used for the main samples. The differences were, first, that no subgroups of schools (high-minority enrollment or private schools) were identified for oversampling (though small schools were still undersampled); second, as explained above, the special supplement to the private school frame was not utilized; and third, the probability of selection for any school in a given age class was capped at 0.5, to ensure that adequate schools remained to be selected for the main sample. In addition, the measure of size used for each school was the estimated number of age eligible students in the school, since for each age class the large majority of students selected were assigned to sessions for which only students of the appropriate age were eligible. The maximum size of the school in which all age-eligible students would be asked to participate was set at 60 for each age class. In most schools having the modal grade, some additional students were selected who were in the modal grade but not age-eligible, so that the maximum sample size of students within a school was about 80 grade- and age-eligible students.

Substitute selections were made for nonparticipating long-term trend sample schools in certain PSUs in a manner similar to the approach used for the main sample schools. As in the case of the main samples, samples of new schools were selected. For age classes 9 and 13, one school was added to the sample in this manner; for age class 17, two schools were added.

Table 3-6 shows the school sample sizes and participation rates for the long-term trend samples for each age class. School participation rates are similar to those seen in equivalent samples in 1990—the fall and winter long-term trend samples and the age 17/grade 11 spring long-term trend sample. As in the main samples, schools selected for both NAEP and NELS were not contacted.

Table 3-6
School Sample Sizes, Refusals, and Substitutes for the Long-term trend Samples

Status	**Age 9/Grade 4 (Winter)**	**Age 13/Grade 8 (Fall)**	**Age 17/Grade 11 (Spring)**	**Total**
Selected, in scope	357	321	267	945
Refusals	45	49	50	144
Participation rate of originally selected schools	87%	85%	81%	85%
1990 participation rate	88%	90%	81%	86%
NELS schools	0	2	40	42
Participating, no eligibles enrolled	9	24	4	37
Substitutes participating	4	3	4	11
Final assessed sample	307	251	217	775

For all three age classes, sessions were assigned to long-term trend sample schools in the following manner. First, the number of sessions per school was established. This was the maximum number of sessions (up to three) that could be administered without creating unduly small session sizes with few eligibles. Thus, in most long-term trend sample schools, three sessions were conducted. However, schools with fewer than 20 eligibles, for example, were asked to conduct only a single session.

The number of session types conducted in the assessment varied by age class. Table 3-7 in the following section shows, among other things, the various long-term trend sample session types conducted for each age class, and the year of the corresponding assessment to which these session types provided a bridge.

The assignment of sessions to schools maximized the number of session types conducted within each PSU. Thus, for example, large schools at age class 9 had a 33.3 percent chance of having one-fifth of their students assigned to the tape-administered mathematics session, and a 66.7 percent chance of not conducting a tape-administered mathematics session at all. Either 20 percent or 40 percent of the students in such schools were assigned to the print-administered reading sessions, in such a way that overall, and as far as possible within each PSU, 30 percent of students were assigned to print-administered mathematics sessions. Just as for the main sample, some schools that were assigned sessions did not participate. As a result, two types of school nonresponse adjustment factors, denoted school and session, were required for the long-term trend samples (see Chapter 10).

This procedure was intended to assure that each session type was assigned to the maximum number of PSUs feasible, given practical constraints in the field. In fact, both print- and tape-administered sessions were given in all 64 PSUs.

3.4 SAMPLING STUDENTS

In the fourth stage of sampling, a consolidated list was prepared for each school of all grade-eligible and age-eligible students for the age class for which the school was selected. A systematic selection of eligible students was made from this list (unless all students were to be assessed) to provide the target sample size. For schools assigned to more than a single session type (the vast majority), students were assigned by Westat district supervisors to one of the various session types using specified procedures. In the long-term trend samples, students assigned to tape-administered sessions who were not age-eligible were dropped from the assessment.

For each age class, separately for the long-term trend and main samples, maxima were established as to the number of students who would be selected for a given school. In those schools that, according to information on the frame, had fewer eligible students than the established maxima, each eligible student enrolled at the school was selected in the sample for one of the sessions assigned to the school. In other schools, a sample of students was drawn, and then students were assigned to sessions as appropriate. For the main samples, the maximum sample sizes were established in terms of the number of grade- plus age-eligible students—150 at age class 9 and 200 at age classes 13 and 17. For the long-term trend samples, the maximum at each age class was 60 age-eligible students or about 80 grade- plus age-eligible students. Note that the number of students actually selected for assessment in a long-term trend sample school generally fell somewhat below 80, because students who were selected for one of the long-term trend tape sessions and were in the modal grade but not age-eligible were subsequently dropped from the sample.

The sample of students to be selected in each school was derived in the following manner, both for main and for long-term trend samples. On the basis of data obtained from the school principal questionnaire (or the sample frame when the principal questionnaire data were not obtained in time) an estimate of the number of eligible students was established for each school. For the main samples, the estimated number of grade- plus age-eligible students was used; for the long-term trend samples, the number of age-eligible students was used. A Session Assignment Form was generated for each school, showing the line numbers (described below) of the students to be selected, indicating the type of session to be taken by each such student. These line numbers were generated using a sampling interval designed to give the appropriate sample size for each school. Thus the overall sampling interval was 1.0 for schools in which all eligible students were to be assessed. The appropriate sampling interval was specified for schools with larger numbers of eligible students, such as to give the appropriate maximum sample size (described above for each age class) in the case that the school had an enrollment of eligible students exactly equal to that predicted.

If the Westat supervisor found that, when applied to the numbered list of eligible students assembled in the field for each school, the line numbers generated gave rise to a sample in excess of 120 percent of the appropriate maximum sample size limit specified above, he or she called Westat's central office. By use of a personal computer, new line numbers based on the actual number of eligible students were generated and relayed to the supervisor. A similar revision to the line numbers was made in the case of a school with a sampling interval in excess of 1.0, and eligible enrollment less than 80 percent of that initially estimated. In this latter case the sample size was increased to the appropriate level. This procedure gave a suitable compromise between control over the sampling rate within each school and operational autonomy and flexibility for Westat field supervisors. Note that in all cases, sampling intervals were generated in Westat's central office, and stored for use in sample weighting. Supervisors were not required to derive or record within-school sampling rates.

Table 3-7 shows the number of students per school who were assessed for each session type. Note that, for the various spiral samples, the number of students assessed per item per school is quite low, even though typically dozens of students were assessed in total in a particular school. Thus the extent of clustering of the sample is in general quite modest, because most sampled schools conducted a few different types of sessions with a moderate number of students in each, and more importantly because the use of BIB-spiraling in the print-administered sessions greatly alleviated the effects of clustering the samples of students within schools.

3.5 EXCLUDED STUDENTS

Some students selected for the sample were deemed unassessable by school authorities because they had limited English language proficiency, were judged as being mildly mentally retarded (educable), or were functionally disabled. In these cases, school staff completed an excluded student questionnaire, listing the reason for exclusion and providing some background information.

Six distinct samples of excluded students were identified. For each age class, there was one sample for the long-term trend assessment and one for the main assessment.

The exclusion criteria for the main samples differed somewhat from those used for the long-term trend samples. The exclusion criteria for the main samples were identical to those used in 1990 and were intended to be somewhat more rigorously defined than those used in the long-term trend samples. (For more details of the exclusion criteria and their implementation, see Chapter 5.) In addition, for age class 17, the excluded students from the long-term trend assessments (with an October-September age definition and modal grade of 11) were drawn from a population different from that of the excluded students in the main assessment (with a calendar-year age definition and modal grade of 12).

For all samples, students were selected for specific sessions, and the school was then asked to identify those to be excluded. Thus only age-eligible students were considered for exclusion from the long-term trend tape-administered sessions, whereas both age- and grade-eligible students were considered for exclusion from the main samples and the print-

Table 3-7

Number of Students per School for Each Session Type

Sample	Sample Code	Session Type	Number of Schools	Mean Number of Students per Session Type per School	Mean Number of Students per Item per School
Age Class 9 Long-term trend	RW-LTTrend	Print booklets 51-56	275	25.7	4.3 - 8.6*
	MS-LTTrend	Tape booklet 91	158	15.1	15.1
	MS-LTTrend	Tape booklet 92	165	15.2	15.2
	MS-LTTrend	Tape booklet 93	161	15.1	15.1
Age Class 9 Main	Math-MainP	Print Mathematics	388	24.3	5.6
	Math-MainT	Tape Mathematics	99	20.7	20.7
	Rdg-MainP	Print Reading	350	24.0	6.0
	Wrt-MainP	Print Writing	375	25.5	5.7
	Math-Calc	Calculator Bridge	102	21.9	21.9
Age Class 13 Long-term trend	RW-LTTrend	Print booklets 51-56	226	24.4	4.1 - 8.1*
	MS-LTTrend	Tape booklet 91	127	15.2	15.2
	MS-LTTrend	Tape booklet 92	130	15.2	15.2
	MS-LTTrend	Tape booklet 93	128	15.7	15.7
Age Class 13 Main	Math-MainP	Print Mathematics	356	28.9	6.7
	Math-MainT	Tape Mathematics	101	23.9	23.9
	Rdg-MainP	Print Reading	412	36.3	5.6
	Wrt-MainP	Print Writing	424	35.2	5.4
Age Class 17 Long-term trend	RW-LTTrend	Print booklets 51-56	256	21.8	3.6 - 7.3*
	MS-LTTrend	Tape booklet 84	153	14.4	14.4
	MS-LTTrend	Tape booklet 85	154	14.0	14.0
Age Class 17 Main	Math-MainP	Print Mathematics	319	29.8	6.9
	Math-MainT	Tape Mathematics	102	20.3	20.3
	Rdg-MainP	Print Reading	396	38.7	3.9
	Wrt-MainP	Print Writing	399	39.3	3.9

---

* This number varied because some item blocks appeared more than once in the set of booklets used for this sample.

administered long-term trend samples. The samples of excluded students for the long-term trend samples were weighted in such a way as to account for this procedure appropriately (see Chapter 10).

Table 3-8 shows the rates of exclusion for each age class for the long-term trend and main samples. For the main samples, for which private school students were oversampled by a factor of three, and constituted about 18 percent of the student sample, exclusion rates are shown for both public and private schools. Overall rates for 1990 (when oversampling of private schools also took place) are shown for comparison. Exclusion rates also appear to have increased slightly over those observed in 1990. Somewhat greater rates of exclusion have occurred in the main than in the long-term trend samples, at least within public schools, but this is somewhat masked by the presence of higher proportions of private school students in the main samples. The most marked effects, however, are the much higher rates of exclusion in public schools than in private, and the higher rates of exclusion at lower grades. The former phenomenon is no doubt a function of the greater prevalence of special education and language minority programs in public schools. The higher exclusion rates at lower ages, which occurred also in 1990 and 1988, result from the greater proportion of students at these grades who are excluded for reasons of limited English proficiency. In certain areas of the United States, fourth-grade public-school students whose native language is Spanish are taught predominantly in Spanish, and in these schools a very high proportion of sampled students are excluded.

Table 3-8
Student Exclusion Rates by Age Class and School Type, Unweighted

Sample	**1992 Exclusion Rate**			**1990 Exclusion Rate**
	Public	**Private**	**Total***	**Total***
Age 9/Grade 4				
Long-term trend	—	—	6.0%	6.1%*
Main	10.5%	1.5%	8.8%	6.5%
Age 13/Grade 8				
Long-term trend	—	—	5.4%	5.5%*
Main	8.7%	1.0%	7.4%	6.1%
Age 17/Grade 11				
Long-term trend	—	—	5.6%	4.4%
Age 17/Grade 12				
Main	6.9%	1.0%	5.9%	4.2%*

---

* Somewhat different exclusion criteria were used for the 1992 main samples than for the 1992 long-term trend samples. Note also that the total rates for 1992 main samples and 1990 samples are based on a relatively greater contribution from private-school students. Private-school students constitute about 18 percent of the sample for the 1992 main samples and the 1990 samples, and about 11 percent of the sample for the 1992 long-term trend samples.

3.6 STUDENT PARTICIPATION RATES

Table 3-9 summarizes the rates of participation of invited students. The set of invited students consists of the selected students, after removing the excluded students and, in the case of long-term trend samples, removing those students selected for tape-administered sessions who were not age-eligible. For a given session, a makeup session was called for when, for various reasons, more than a tolerable number of invited students failed to attend the originally scheduled session to which they were invited. The participation rates given in the table express the number finally assessed as a percentage of those initially invited in the participating schools. Participation rates are shown for the main and long-term trend samples and for public and private schools separately in the case of the main samples. Overall participation rates are also shown for comparable samples from the 1990 NAEP assessment. The table shows that student participation rates in 1992 are very similar to those experienced in 1990. At all age classes, the participation rate of private school students exceeds that of public school students, with the difference, both relative and absolute, increasing with age class. This is in contrast with the levels of school participation, which are higher for public schools.

Table 3-9
Student Participation Rates by Age Class and School Type, Unweighted

Samples	1992 Public		1992 Private		1992 Combined		
	Number Invited	Participation Rate	Number Invited	Participation Rate	Number Invited	Participation Rate	1990 Participation Rate
Age 9/Grade 4							
Long-term trend	—	—	—	—	15,321	94.0%	92.4%
Main	27,291	93.0%	6,607	95.4%	33,898	93.4%	92.9%*
Age 13/Grade 8							
Long-term trend	—	—	—	—	12,574	90.8%	90.4%
Main	39,465	87.5%	8,499	94.6%	47,964	88.8%	89.1%*
Age 17/Grade 11							
Long-term trend	—	—	—	—	11,985	82.8%	81.2%
Age 17/Grade 12							
Main	43,340	79.0%	9,346	89.2%	52,686	80.8%	81.3%*

---

* The total rates for the 1992 main samples and the 1990 samples are based on a relatively greater contribution of private-school students than the 1992 long-term trend samples. Private-school students constitute about 18 percent of the invited students for the 1992 main samples and the 1990 samples, and about 11 percent of the invited students for the 1992 long-term trend samples.

3.7 OVERALL STUDENT PARTICIPATION RATES

The combined impact of school nonparticipation and student absenteeism from sessions within participating schools is summarized in Table 3-10. The table shows the percentages of students assessed, from among those who would have been assessed if all initially selected schools had participated, and if all invited students had attended either an initial or make-up session. The results show that, consistent with earlier rounds of NAEP, the overall level of participation decreases substantially with the increase in age and grade of the students.

Table 3-10
Overall Participation Rates (School and Student Combined) by Age Class, Unweighted

1992 Samples	Age Class 9	Age Class 13	Age Class 17	Overall
Main Samples				
School participation	86.4%	85.3%	81.5%	84.5%
Student participation	93.4%	88.8%	80.8%	86.8%
Overall student participation	80.7%	75.7%	65.9%	73.3%
Number of participating students	31,672	42,591	42,556	116,819
Long-term Trend Samples				
School participation	87.4%	84.7%	81.3%	84.8%
Student participation	94.0%	90.8%	82.8%	89.6%
Overall student participation	82.2%	76.9%	67.3%	75.9%
Number of participating students	14,397	11,423	9,928	35,748
Overall				
School participation	86.8%	85.1%	81.4%	84.6%
Student participation	93.6%	89.2%	81.2%	87.5%
Overall student participation	81.2%	75.9%	66.1%	74.0%
Number of participating students	46,069	54,014	52,484	152,567

Until now, only unweighted participation rates by age class and school type have been presented. However, analysis is typically performed separately by age class and session type, and NCES and ETS quote rules regarding acceptable potentials for bias in terms of weighted participation rates. Therefore, Tables 3-11 and 3-12 show weighted participation rates by age class and session type for the main and long-term trend samples, respectively. Note that these participation rates are similar for different session types in the same age class. They are also similar, in general, to the unweighted rates.

The procedures for substituting for nonparticipating schools or imputing for them and the procedures for imputing for absent students were designed (so far as feasible) to reduce the biases resulting from school and student nonparticipation. These procedures are discussed in Chapter 10.

Table 3-11

Weighted Participation Rates by Age Class and Session Type, Main Samples

Participation	Calculator Bridge	Mathematics Print	Reading Print	Mathematics Tape	Writing Print
Age Class 9					
School participation	88.1%	85.9%	86.9%	87.9%	85.9%
Student participation	94.4%	93.5%	93.5%	92.7%	93.3%
Overall participation	83.1%	80.3%	81.3%	81.5%	80.1%
Age Class 13					
School participation	-	85.6%	83.7%	85.1%	86.1%
Student participation	-	89.5%	89.1%	89.7%	88.9%
Overall participation	-	76.5%	74.6%	76.3%	76.6%
Age Class 17					
School participation	-	82.0%	81.2%	88.5%	82.0%
Student participation	-	81.4%	81.1%	80.2%	80.7%
Overall participation	-	66.8%	65.8%	71.0%	66.2%

Table 3-12

Weighted Participation Rates by Age Class and Session Type, Long-term Trend Samples

Participation	Reading/Writing Print	Math/Science Tape
Age Class 9		
School participation	87.0%	87.8%
Student participation	93.8%	94.4%
Overall participation	81.5%	82.9%
Age Class 13		
School participation	85.3%	85.6%
Student participation	90.8%	90.9%
Overall participation	77.4%	77.8%
Age Class 17		
School participation	80.9%	81.0%
Student participation	83.3%	82.3%
Overall participation	67.4%	66.6%

3.8 SAMPLING TEACHERS

The teacher questionnaire was administered to mathematics, reading, and writing teachers of fourth-grade and eighth-grade students assessed in mathematics, fourth-grade students assessed in reading, and eighth-grade students assessed in writing. Teachers were selected if they taught the student the subject in which the student was assessed. The purpose of drawing these samples was not to estimate the attributes of the teacher population, but to estimate the number (proportion) of students whose teachers had various attributes and to correlate student characteristics and performance with the characteristics of their teachers.

The selected teachers were asked to complete a questionnaire concerning themselves and their teaching practices, with specific references to each individual class period containing a student included in the main assessment.

Chapter 4

ASSESSMENT INSTRUMENTS

Debra L. Kline, Stephen Lazer, and Jay R. Campbell

Educational Testing Service

In the 1992 assessment, five types of instruments were used to collect data about students, teachers, and schools. Each assessed student received an **assessment booklet** containing both cognitive and background questions. An **excluded student questionnaire** was completed by school officials for each sampled student who was deemed unable to take part in the assessment. The mathematics teachers of fourth- and eighth-grade students participating in the mathematics assessment, the reading teachers of fourth-grade students participating in the reading assessment, and the writing teachers of eighth-grade students in the writing assessment were asked to complete a **teacher questionnaire**. A **school characteristics and policies questionnaire** was distributed to each participating school. A **principal questionnaire** was completed prior to the assessment for each school selected for the sample.

In addition, as part of a special study of school-based writing, NAEP collected writing papers from random samples of students at grades 4 and 8.

This chapter begins with a discussion of the characteristics of the student booklets used for the 1992 main and long-term trend assessments and how the booklets were assembled. A set of tables presents in detail the contents of each booklet and item block. Sections 4.3 through 4.7 provide an overview of the contents of the student background questionnaires, teacher questionnaires, excluded student questionnaires, school questionnaires, and principal questionnaires. Section 4.8 describes the special study that was part of the writing assessment.

4.1 STUDENT BOOKLETS—MAIN ASSESSMENTS

4.1.1 Reading

Each student assessed in reading received a booklet containing a set of general background questions, reading passages and content questions, a set of subject-specific background questions, and a set of questions about his or her motivation and familiarity with the assessment materials. The passages and content questions were assembled into sections or blocks, each containing either one or two passages and the corresponding questions. Students were given either two 25-minute blocks or one 50-minute block. The overall assessment time for each student was approximately 63 minutes.

The assembly of reading blocks into booklets and their subsequent assignment to sampled students was determined by a *partially balanced incomplete block* (PBIB) design with *spiraled* administration. These blocks were then assembled into booklets containing two 5-minute background sections, one 3-minute background section, and two 25-minute blocks of reading passages and items according to a partially balanced incomplete block design.[1]

The reading blocks were assigned to booklets in such a way that every 25-minute block within a given purpose for reading (either Reading for Literary Experience, Reading to Gain Information, or Reading to Perform a Task) was paired with every other block measuring the same purpose but was only paired with one block measuring another purpose for reading. At all grades, every 25-minute block appears in four booklets. At grade 4, each block is paired three times with another block that measures the same purpose for reading. A given block is also paired once with a block measuring the other purpose for reading. At grades 8 and 12, every 25-minute block is paired once with each of the two other blocks that measure the same purpose for reading. A given block is also paired with one block measuring each of the other two purposes for reading. This is the *partially balanced* part of the balanced incomplete block design. Every 50-minute block appears only in one booklet.

The PBIB design for the 1992 national reading assessment was *focused* by subject area, so that students received booklets containing only blocks of reading questions (not mathematics or writing). The *focused*-PBIB design also balances the order of presentation of the 25-minute blocks of items—every 25-minute block appears as the first cognitive block in two booklets and as the second cognitive block in two other booklets. This design allows for some control of context and fatigue effects.

The design used in 1992 required that eight blocks of grade 4 reading items be assembled into sixteen booklets. At grades 8 and 12 the 25-minute blocks were assembled into 18 booklets. At grade 8, there were two additional booklets containing 50-minute blocks; at grade 12 there were three of these booklets. Once assembled, the assessment booklets were then *spiraled* and bundled. Spiraling involves interweaving the booklets in a systematic sequence so that each booklet appears an appropriate number of times in the sample. The bundles were designed so that each booklet would appear equally often in each position in a bundle.

The final step in the PBIB-spiraling procedure was the assigning of the booklets to the assessed students. The students within an assessment session were assigned booklets in the order in which the booklets were bundled. Thus, most students in an assessment session received different booklets. In the assessment design, representative and randomly equivalent samples of about 2,500 students responded to each item at a given age/grade level.

Tables 4-1, 4-3, and 4-5 provide the composition and number of blocks administered in the 1992 reading assessment. Tables 4-2, 4-4, and 4-6 give details of the item blocks used in the main reading assessment, including the number of cognitive and constructed-response items in each block and the booklets in which each block appeared.

---

[1] Booklets containing 50-minute blocks are included in the spiraled administration, but cannot be assembled in the partially balanced incomplete block fashion.

Table 4-1
Main Sample Booklet Contents and Number of Booklets Administered
Age 9/grade 4, Reading

Subject Area	Booklet Number	Common Background Block	Cognitive Blocks		Subject Area Background Block	Motivation Background Block	Number of Booklets Administered
Reading	30	CA	R4	R3	R2	RX	525
	31	CA	R3	R5	R2	RX	534
	32	CA	R5	R9	R2	RX	536
	33	CA	R9	R4	R2	RX	526
	34	CA	R4	R5	R2	RX	523
	35	CA	R3	R9	R2	RX	517
	36	CA	R6	R10	R2	RX	534
	37	CA	R10	R7	R2	RX	522
	38	CA	R7	R8	R2	RX	517
	39	CA	R8	R6	R2	RX	528
	40	CA	R6	R7	R2	RX	517
	41	CA	R10	R8	R2	RX	528
	42	CA	R7	R4	R2	RX	533
	43	CA	R8	R3	R2	RX	531
	44	CA	R5	R6	R2	RX	518
	45	CA	R9	R10	R2	RX	527

Table 4-2
Cognitive and Noncognitive Block Information
Age 9/grade 4, Reading

Block	Type*	Total Number of Items	Number of Multiple-choice Items	Number of Constructed-response Items	Booklets Containing Block
CA	Common Background	20	20	0	30 - 45
R2	Reading Background	14	14	0	30 - 45
RX	Reading Motivation	5	5	0	30 - 45
R3	Reading Cognitive (L)	11	6	5	30, 31, 35, 43
R4	Reading Cognitive (L)	12	5	7	30, 33, 34, 42
R5	Reading Cognitive (L)	11	7	4	31, 32, 34, 44
R6	Reading Cognitive (I)	10	5	5	36, 39, 40, 44
R7	Reading Cognitive (I)	10	4	6	37, 38, 40, 42
R8	Reading Cognitive (I)	10	5	5	38, 39, 41, 43
R9	Reading Cognitive (L)	9	4	5	32, 33, 35, 45
R10	Reading Cognitive (I)	12	6	6	36, 37, 41, 45

---

* At grade 4, all reading cognitive blocks are 25 minutes in length. The letter "L" denotes a block designed to measure reading for literary experience, and "I" indicates reading to gain information.

BEST COPY AVAILABLE

Table 4-3
Main Sample Booklet Contents and Number of Booklets Administered
Age 13/grade 8, Reading

Subject Area	Booklet Number	Common Background Block	Cognitive Blocks*		Subject Area Background Block	Motivation Background Block	Number of Booklets Administered
Reading	30	CA	R3	R4	RB	RX	561
	31	CA	R4	R5	RB	RX	582
	32	CA	R5	R3	RB	RX	569
	33	CA	R6	R8	RB	RX	566
	34	CA	R8	R7	RB	RX	575
	35	CA	R7	R6	RB	RX	587
	36	CA	R10	R9	RB	RX	574
	37	CA	R9	R11	RB	RX	573
	38	CA	R11	R10	RB	RX	572
	39	CA	R3	R8	RB	RX	584
	40	CA	R7	R4	RB	RX	597
	41	CA	R5	R6	RB	RX	577
	42	CA	R6	R9	RB	RX	558
	43	CA	R8	R11	RB	RX	560
	44	CA	R10	R7	RB	RX	567
	45	CA	R4	R10	RB	RX	595
	46	CA	R9	R5	RB	RX	556
	47	CA	R11	R3	RB	RX	565
	48	CA	R12		RB	RX	2297
	49	CA	R13		RB	RX	2327

---

* Block R9 requires a bus schedule; block R12 requires the NAEP Reader; blocks R12, and R13 are 50-minute blocks.

Table 4-4
Cognitive and Noncognitive Block Information
Age 13/grade 8, Reading

Block	Type*	Total Number of Items	Number of Multiple-choice Items	Number of Constructed-response Items	Booklets Containing Block
CA	Common Background	22	22	0	30 - 49
R2	Reading Background	24	24	0	30 - 49
RX	Reading Motivation	5	5	0	30 - 49
R3	Reading Cognitive (L)	11	4	7	30, 32, 39, 47
R4	Reading Cognitive (L)	13	8	5	30, 31, 40, 45
R5	Reading Cognitive (L)	11	7	4	31, 32, 41, 46
R6	Reading Cognitive (I)	12	5	7	33, 35, 41, 42
R7	Reading Cognitive (I)	13	6	7	34, 35, 40, 44
R8	Reading Cognitive (I)	14	7	7	33, 34, 39, 43
R9	Reading Cognitive (T)	12	7	5	36, 37, 42, 46
R10	Reading Cognitive (T)	12	4	8	36, 38, 44, 45
R11	Reading Cognitive (T)	12	3	9	37, 38, 43, 47
R12	Reading Cognitive (L)	12	0	12	48
R13	Reading Cognitive (I)	13	5	8	49

---

* All reading cognitive blocks are 25 minutes long except R12 and R13, which are 50 minutes long. The letter "L" denotes a block designed to measure reading for literary experience, "I" indicates reading to gain information, and "T" denotes reading to perform a task.

Table 4-5
Main Sample Booklet Contents and Number of Booklets Administered
Age 17/grade 12, Reading

Subject Area	Booklet Number	Common Background Block	Cognitive Blocks*		Subject Area Background Block	Motivation Background Block	Number of Booklets Administered
Reading	30	CA	R3	R4	RB	RX	518
	31	CA	R4	R5	RB	RX	517
	32	CA	R5	R3	RB	RX	510
	33	CA	R6	R7	RB	RX	507
	34	CA	R7	R8	RB	RX	502
	35	CA	R8	R6	RB	RX	494
	36	CA	R10	R9	RB	RX	497
	37	CA	R9	R11	RB	RX	500
	38	CA	R11	R10	RB	RX	498
	39	CA	R3	R7	RB	RX	506
	40	CA	R8	R4	RB	RX	506
	41	CA	R5	R6	RB	RX	518
	42	CA	R6	R9	RB	RX	529
	43	CA	R7	R11	RB	RX	517
	44	CA	R9	R8	RB	RX	528
	45	CA	R4	R10	RB	RX	530
	46	CA	R9	R5	RB	RX	509
	47	CA	R11	R3	RB	RX	515
	48	CA	R12		RB	RX	2024
	49	CA	R13		RB	RX	2017
	50	CA	R14		RB	RX	2073

---

* Block R9 requires a bus schedule; block R11 requires a tax form; block R12 requires the NAEP Reader; blocks R12, R13, and R14 are 50-minute blocks.

Table 4-6
Cognitive and Noncognitive Block Information
Age 17/grade 12, Reading

Block	Type*	Total Number of Items	Number of Multiple-choice Items	Number of Constructed-response Items	Booklets Containing Block
CA	Common Background	29	29	0	30 - 50
R2	Reading Background	24	24	0	30 - 50
RX	Reading Motivation	5	5	0	30 - 50
R3	Reading Cognitive (L)	11	4	7	30, 32, 39, 47
R4	Reading Cognitive (L)	9	3	6	30, 31, 40, 45
R5	Reading Cognitive (L)	12	5	7	31, 32, 41, 46
R6	Reading Cognitive (I)	12	5	7	33, 35, 41, 42
R7	Reading Cognitive (I)	12	5	7	33, 34, 39, 43
R8	Reading Cognitive (I)	10	3	7	34, 35, 40, 44
R9	Reading Cognitive (T)	12	7	5	36, 37, 42, 46
R10	Reading Cognitive (T)	12	4	8	36, 38, 44, 45
R11	Reading Cognitive (T)	15	7	8	37, 38, 43, 47
R12	Reading Cognitive (L)	12	0	12	48
R13	Reading Cognitive (I)	16	10	6	49
R14	Reading Cognitive (I)	12	7	5	50

---

* All reading cognitive blocks are 25 minutes long except R12, R13, and R14, which are 50 minutes long The letter "L" denotes a block designed to measure reading for literary experience, "I" indicates reading to gain information, and "T" denotes reading to perform a task.

4.1.2 Mathematics

Each student in the main mathematics assessment received a booklet containing a set of general background questions, a set of subject-specific background questions, three 15-minute segments or blocks of cognitive items, and a set of questions about his or her motivation and familiarity with the assessment material. At each grade level, the mathematics assessment included 16 different blocks of multiple-choice and constructed-response content questions. Students received different blocks of cognitive items in their booklets according to a careful plan. The 1992 assessment was based on an adaptation of matrix sampling called balanced incomplete block (BIB) spiraling—a design that enables broad coverage of mathematics content while minimizing the burden for any one student. The balanced incomplete block part of the design assigns blocks of items to booklets and each pair of blocks appears together in at least one booklet. It also ensures that blocks appear in each possible position in an examination booklet, to balance context and fatigue effects. The spiraling part of the method cycles the booklets for administration, so that typically only a few students in any assessment session receive the same booklet.

Thirteen of the 16 blocks were assembled in accordance with this design, whereby the 13 blocks were presented in 26 booklets. Each block appeared in exactly six booklets, and each block appeared with every other block in at least one booklet. Students at grades 4 and 8 were given calculators to use with three of the 13 blocks and were trained in their use prior to the assessment. Students at grade 12 were given calculators to use with four of the 13 blocks. At the fourth grade, students were provided with four-function calculators and at grades 8 and 12, they were provided with scientific calculators. For another of the blocks, fourth-grade students were provided with a ruler, and eighth- and twelfth-grade students with a protractor/ruler. For still another of the blocks, at all three grades, students were given geometric shapes (manipulatives) to provide a concrete basis for determining their answers.

For the national assessment, the three remaining blocks at each grade used a paced-audiotape format to measure students' estimation skills and to move students through some word problems at grade 4 as well as material measuring data analysis, probability, and statistics and pre-algebra at the two upper grades. For the estimation block, the pacing method curtails time for computations and in more complex problem situations, it facilitates instances where students might have difficulty in reading the question or might spend too little or too much time on particular questions. The three blocks accompanied by the audiotape were assembled into one booklet at each grade. Of the 16 blocks, five were carried forward from 1990 to use in measuring trends across time, including four of the 13 BIB-spiraled blocks and the estimation block.

In addition to the three blocks of cognitive questions, each participating student answered two five-minute sets of background questions, one general and one on instructional, behavioral, and attitudinal variables related to mathematics. A one-minute questionnaire was given to students at the end of each booklet to determine students' motivation in completing the assessment and their familiarity with assessment tasks.

Tables 4-7, 4-8, and 4-9 give the composition and number of mathematics booklets administered in the 1992 main assessment. Tables 4-10, 4-11, and 4-12 provide the total number

Table 4-7

Main Sample Booklet Contents and Number of Booklets Administered
Age 9/grade 4, Mathematics

Subject Area	Booklet Number	Common Background Block	Subject Area Background Block	Cognitive Blocks*				Motivation Background Block	Number of Booklets Administered
Mathematics	1	CA	M2	M3	M4	M7		MX	354
	2	CA	M2	M4	M5	M8		MX	346
	3	CA	M2	M5	M6	M9		MX	358
	4	CA	M2	M6	M7	M10		MX	359
	5	CA	M2	M7	M8	M11		MX	356
	6	CA	M2	M8	M9	M12		MX	367
	7	CA	M2	M9	M10	M13		MX	360
	8	CA	M2	M10	M11	M14		MX	368
	9	CA	M2	M11	M12	M15		MX	375
	10	CA	M2	M12	M13	M3		MX	377
	11	CA	M2	M13	M14	M4		MX	374
	12	CA	M2	M14	M15	M5		MX	380
	13	CA	M2	M15	M3	M6		MX	370
	14	CA	M2	M3	M5	M10		MX	382
	15	CA	M2	M4	M6	M11		MX	361
	16	CA	M2	M5	M7	M12		MX	356
	17	CA	M2	M6	M8	M13		MX	367
	18	CA	M2	M7	M9	M14		MX	362
	19	CA	M2	M8	M10	M15		MX	365
	20	CA	M2	M9	M11	M3		MX	350
	21	CA	M2	M10	M12	M4		MX	357
	22	CA	M2	M11	M13	M5		MX	347
	23	CA	M2	M12	M14	M6		MX	355
	24	CA	M2	M13	M15	M7		MX	361
	25	CA	M2	M14	M3	M8		MX	357
	26	CA	M2	M15	M4	M9		MX	350
Estimation & Complex Problem Solving	27	CA	M2	M4	M16	M17	M18	MX	2054
Calculator Bridge	28	CA	M2	M14	M3	M18		MX	2236

---

* Block M5 requires a ruler; blocks M8, M12, and M14 require a calculator; block MJ requires geometric shapes; blocks M7, M9, M13, M14, and M15 contain extended constructed-response items; block M16 contains estimation items; blocks M17 and M18 contain complex problem-solving items.

Table 4-8

Main Sample Booklet Contents and Number of Booklets Administered
Age 13/grade 8, Mathematics

Subject Area	Booklet Number	Common Background Block	Subject Area Background Block	Cognitive Blocks*	Motivation Background Block	Number of Booklets Administered
Mathematics	1	CA	M2	M3 M4 M7	MX	395
	2	CA	M2	M4 M5 M8	MX	382
	3	CA	M2	M5 M6 M9	MX	402
	4	CA	M2	M6 M7 M10	MX	402
	5	CA	M2	M7 M8 M11	MX	394
	6	CA	M2	M8 M9 M12	MX	390
	7	CA	M2	M9 M10 M13	MX	390
	8	CA	M2	M10 M11 M14	MX	394
	9	CA	M2	M11 M12 M15	MX	394
	10	CA	M2	M12 M13 M3	MX	386
	11	CA	M2	M13 M14 M4	MX	403
	12	CA	M2	M14 M15 M5	MX	399
	13	CA	M2	M15 M3 M6	MX	394
	14	CA	M2	M3 M5 M10	MX	405
	15	CA	M2	M4 M6 M11	MX	389
	16	CA	M2	M5 M7 M12	MX	396
	17	CA	M2	M6 M8 M13	MX	397
	18	CA	M2	M7 M9 M14	MX	393
	19	CA	M2	M8 M10 M15	MX	385
	20	CA	M2	M9 M11 M3	MX	388
	21	CA	M2	M10 M12 M4	MX	401
	22	CA	M2	M11 M13 M5	MX	401
	23	CA	M2	M12 M14 M6	MX	398
	24	CA	M2	M13 M15 M7	MX	414
	25	CA	M2	M14 M3 M8	MX	399
	26	CA	M2	M15 M4 M9	MX	400
Estimation & Complex Problem Solving	27	CA	M2	M4 M16 M17 M18	MX	2416

---

* Block M5 requires a protractor/ruler; blocks M8, M12, and M14 require a calculator; block M10 requires geometric shapes; blocks M3, M7, M9, M12, M13, and M14 contain extended constructed-response items; block M16 contains estimation items; blocks M17 and M18 contain complex problem-solving items.

Table 4-9

Main Sample Booklet Contents and Number of Booklets Administered
Age 17/grade 12, Mathematics

Subject Area	**Booklet Number**	**Common Background Block**	**Subject Area Background Block**	**Cognitive Blocks***				**Motivation Background Block**	**Number of Booklets Administered**
Mathematics	1	CA	M2	M3	M4	M7		MX	370
	2	CA	M2	M4	M5	M8		MX	366
	3	CA	M2	M5	M6	M9		MX	366
	4	CA	M2	M6	M7	M10		MX	382
	5	CA	M2	M7	M8	M11		MX	377
	6	CA	M2	M8	M9	M12		MX	364
	7	CA	M2	M9	M10	M13		MX	355
	8	CA	M2	M10	M11	M14		MX	363
	9	CA	M2	M11	M12	M15		MX	383
	10	CA	M2	M12	M13	M3		MX	372
	11	CA	M2	M13	M14	M4		MX	363
	12	CA	M2	M14	M15	M5		MX	363
	13	CA	M2	M15	M3	M6		MX	353
	14	CA	M2	M3	M5	M10		MX	342
	15	CA	M2	M4	M6	M11		MX	360
	16	CA	M2	M5	M7	M12		MX	362
	17	CA	M2	M6	M8	M13		MX	350
	18	CA	M2	M7	M9	M14		MX	367
	19	CA	M2	M8	M10	M15		MX	364
	20	CA	M2	M9	M11	M3		MX	362
	21	CA	M2	M10	M12	M4		MX	366
	22	CA	M2	M11	M13	M5		MX	378
	23	CA	M2	M12	M14	M6		MX	364
	24	CA	M2	M13	M15	M7		MX	370
	25	CA	M2	M14	M3	M8		MX	368
	26	CA	M2	M15	M4	M9		MX	369
Estimation & Complex Problem Solving	27	CA	M2	M4	M16	M17	M18	MX	2074

---

* Block M13 requires a protractor/ruler; blocks M7, M8, M12, and M14 require a calculator; block M10 requires geometric shapes; blocks M9, M10, M12, M13, M14, and M15 contain extended constructed-response items; block M16 contains estimation items; and blocks M17 and M18 contain complex problem-solving items.

Table 4-10

Cognitive and Noncognitive Block Composition
Age 9/grade 4, Mathematics

Block	Type	Total Number of Items	Number of Multiple-Choice Items	Number of Constructed-response Items		Booklets* Containing Block
				Short	Extended	
B1	Common Background	20	20	0	0	1 - 28
M2	Mathematics Background	18	18	0	0	1 - 28
MB	Motivation Background	5	5	0	0	1 - 28
M3	Mathematics Cognitive	13	9	4	0	1,10,13,14,20,25,28
M4	Mathematics Cognitive (Trend)	14	14	0	0	1, 2,11,15,21,26,27
M5	Mathematics Cognitive (Trend/Ruler)	17	13	4	0	2, 3,12,14,16,22
M6	Mathematics Cognitive (Trend)	11	0	11	0	3, 4,13,15,17,23
M7	Mathematics Cognitive	10	6	3	1	1, 4, 5,16,18,24
M8	Mathematics Cognitive (Trend/Calculator)	15	14	1	0	2, 5, 6,17,19,25,28
M9	Mathematics Cognitive	12	9	2	1	3, 6, 7,18,20,26
M10	Mathematics Cognitive (Manipulatives)	6	0	6	0	4, 7, 8,14,19,21
M11	Mathematics Cognitive	16	11	5	0	5, 8, 9,15,20,22
M12	Mathematics Cognitive (Calculator)	12	5	7	0	6, 9,10,16,21,23
M13	Mathematics Cognitive	12	6	5	1	7,10,11,17,22,24
M14	Mathematics Cognitive (Calculator)	10	6	3	1	8,11,12,18,23,25,28
M15	Mathematics Cognitive	10	6	3	1	9,12,13,19,24,26
M16	Mathematics Cognitive (Estimation)	20	20	0	0	27,29
M17	Mathematics Cognitive (Pre-algebra)	7	1	6	0	27
M18	Mathematics Cognitive (Word problems)	10	4	6	0	27

---

* Booklets M1 to M26 represent the main BIB-spiral assessment; booklet 27 was used for the special paced-tape blocks and booklet 29 was used for the paced-tape sections of the Trial State Assessment.

Table 4-11

Cognitive and Noncognitive Block Information
Age 13/grade 8, Mathematics

Block	Type	Total Number of Items	Number of Multiple-Choice Items	Number of Constructed-response Items		Booklets* Containing Block
				Short	Extended	
B1	Common Background	22	22	0	0	1 - 27
M2	Mathematics Background	23	23	0	0	1 - 27
MB	Motivation Background	5	5	0	0	1 - 27
M3	Mathematics Cognitive	13	9	3	1	1,10,13,14,20,25
M4	Mathematics Cognitive (Trend)	21	21	0	0	1, 2,11,15,21,26,27
M5	Mathematics Cognitive (Trend/Ruler)	21	16	5	0	2, 3,12,14,16,22
M6	Mathematics Cognitive (Trend)	16	0	16	0	3, 4,13,15,17,23
M7	Mathematics Cognitive	13	7	5	1	1, 4, 5,16,18,24
M8	Mathematics Cognitive (Trend/Calculator)	18	16	2	0	2, 5, 6,17,19,25
M9	Mathematics Cognitive	9	5	3	1	3, 6, 7,18,20,26
M10	Mathematics Cognitive (Manipulatives)	7	0	7	0	4, 7, 8,14,19,21
M11	Mathematics Cognitive	19	13	6	0	5, 8, 9,15,20,22
M12	Mathematics Cognitive (Calculator)	9	6	2	1	6, 9,10,16,21,23
M13	Mathematics Cognitive	11	6	4	1	7,10,11,17,22,24
M14	Mathematics Cognitive (Calculator)	9	6	2	1	8,11,12,18,23,25
M15	Mathematics Cognitive	17	13	4	0	9,12,13,19,24,26
M16	Mathematics Cognitive (Estimation)	22	22	0	0	27,29
M17	Mathematics Cognitive (Pre-algebra)	11	4	7	0	27
M18	Mathematics Cognitive (Data analysis)	6	4	2	0	27

---

* Booklets M1 to M26 represent the main BIB-spiral assessment; booklet 27 was used for the special paced-tape blocks and booklet 29 was used for the paced-tape sections of the Trial State Assessment.

Table 4-12

Cognitive and Noncognitive Block Information
Age 17/grade 12, Mathematics

Block	Type	Total Number of Items	Number of Multiple-Choice Items	Number of Constructed-response Items		Booklets* Containing Block
				Short	Extended	
B1	Common Background	29	29	0	0	1 - 27
M2	Mathematics Background	34	34	0	0	1 - 27
MB	Motivation Background	5	5	0	0	1 - 27
M3	Mathematics Cognitive	14	10	4	0	1,10,13,14,20,25
M4	Mathematics Cognitive (Trend)	22	22	0	0	1, 2,11,15,21,26,27
M5	Mathematics Cognitive (Trend)	20	17	3	0	2, 3,12,14,16,22
M6	Mathematics Cognitive (Trend)	17	0	17	0	3, 4,13,15,17,23
M7	Mathematics Cognitive	13	7	6	0	1, 4, 5,16,18,24
M8	Mathematics Cognitive (Trend/Calculator)	21	17	4	0	2, 5, 6,17,19,25
M9	Mathematics Cognitive	9	6	2	1	3, 6, 7,18,20,26
M10	Mathematics Cognitive (Manipulatives)	10	3	6	1	4, 7, 8,14,19,21
M11	Mathematics Cognitive	14	11	3	0	5, 8, 9,15,20,22
M12	Mathematics Cognitive (Calculator)	9	5	3	1	6, 9,10,16,21,23
M13	Mathematics Cognitive (Prot/Ruler)	9	3	5	1	7,10,11,17,22,24
M14	Mathematics Cognitive (Calculator)	10	6	3	1	8,11,12,18,23,25
M15	Mathematics Cognitive	11	8	2	1	9,12,13,19,24,26
M16	Mathematics Cognitive (Estimation)	22	22	0	0	27
M17	Mathematics Cognitive (Pre-algebra)	12	5	7	0	27
M18	Mathematics Cognitive (Data analysis)	12	7	5	0	27

---

* Booklets M1 to M26 represent the main BIB-spiraled assessment; booklet 27 was used for the special paced-tape blocks.

of booklets, cognitive blocks, and noncognitive blocks used for the program at each grade level. The tables also include details of the BIB design used in the mathematics assessment.

4.1.3 Writing

At each grade, three of the 25-minute writing tasks measured each of the purposes for writing specified in the framework. At grade 8, one of the 50-minute exercises was in the area of informative writing, and the other in the area of narrative writing. At grade 12, one 50-minute task was devoted to each purpose for writing.

Each student received an assessment booklet containing a set of general background questions, writing tasks (or in some cases a single task), a set of subject-specific background questions, and a set of questions about his or her motivation and familiarity with the assessment materials. Students were given either two 25-minute exercises or one 50-minute exercise. In addition to the one or two cognitive blocks, students received three sections of background questions. The overall assessment time for each student was approximately 63 minutes,

The assembly of writing blocks into booklets and their subsequent assignment to sampled students was determined by a *partially balanced incomplete block* (PBIB) design with *spiraled* administration. The blocks were assigned to booklets in such a way that every 25-minute block within a given purpose for writing was paired with every other block measuring the same purpose but was only paired with one block measuring another purpose for writing. At all grades, every 25-minute block appears in four booklets. This is the *partially balanced* part of the balanced incomplete block design. Every 50-minute block appears only in one booklet (although booklets containing the 50-minute blocks are included in the main BIB-spiraled assessment, they cannot be assembled in the PBIB fashion.)

Each 25-minute block was paired with every other writing block assessing the same purpose for writing but not with all the blocks assessing the other purpose for writing. The *focused*-PBIB design also balances the order of presentation of the 25-minute blocks—every 25-minute block appears as the first cognitive block in two booklets and as the second cognitive block in two other booklets. This design allows for some control of context and fatigue effects.

At each grade in the assessment, the 25-minute tasks were assembled into 18 booklets. At grade 8 there were two additional booklets containing 50-minute tasks, while at grade 12 there were three such booklets. The assessment booklets were then spiraled and bundled. Spiraling involves interweaving the booklets in a systematic sequence so that each booklet appears an appropriate number of times in the sample. The bundles were designed so that each booklet would appear equally often in a position in a bundle.

As in the other subjects, the final step in the BIB-spiraling procedure was the assigning of booklets to the assessed students. The students in the assessment session were assigned booklets in the order in which the booklets were bundled. Thus, most students in an assessment session received different booklets. In the design, representative and randomly equivalent samples of about 2,500 students responded to each exercise at a given age/grade level.

Tables 4-13, 4-15, and 4-17 detail the composition and number of booklets administered in the 1992 writing assessment, and indicate which writing prompts measured specific purposes for writing. Tables 4-14, 4-16, and 4-18 provide the total number of booklets, cognitive blocks, and noncognitive blocks used for the program at each grade level. These tables also provide the details of the focused-PBIB design used in the writing assessment.

4.2 STUDENT BOOKLETS—LONG-TERM TREND ASSESSMENTS

There were several long-term trend samples in the 1992 assessment (see Chapter 1), each of which required the use of special booklets. Tables 4-19, 4-20, and 4-21 summarize the contents of each trend assessment booklet and show how many of each booklet were administered. Tables 4-22, 4-23, and 4-24 give details of the item blocks used in the long-term trend assessments, including the number of cognitive and constructed-response items in each block and the booklets in which each block appeared.

Reading and Writing Long-term Trend. Six booklets (numbered 51 to 56) containing reading and writing items were administered to each age class. These booklets were identical to booklets used in previous assessments of reading and writing and were BIB-spiraled for administration. Each booklet consisted of a common background block (BZ) and three cognitive blocks (at least one reading block and at least one writing block). In addition to cognitive items, the cognitive blocks also contained subject-related background questions.

Mathematics and Science Long-term Trend. Three booklets (91, 92, and 93) at ages 9 and 13 and two booklets (84 and 85) at age 17, containing mathematics and science items, were identical to those used in previous assessments to measure trends. Each booklet contained a common background block (C1 or BZ) and three cognitive blocks. At ages 9 and 13, these booklets contained one reading block (R1, R2, or R3), one mathematics block (M1, M2, or M3) and one science block (S1, S2, or S3). At age 17, each booklet contained at least one mathematics block (M1 to M3) and at least one science block (S1 - S3). Mathematics block M3 contained items that required the use of a calculator. All cognitive blocks also contained subject-related background questions.

4.3 MAIN ASSESSMENT STUDENT QUESTIONNAIRES

Each booklet in the main assessment included three student background questionnaires. The first, consisting of general background questions, included questions about race/ethnicity, mother's and father's level of education, reading materials in the home, homework, attendance, academic expectations, and which parents lived at home. The second, consisting of subject-area background questions, included questions about instructional activities, courses taken, use of specialized resources such as calculators in mathematics class, and views on the utility and value of the subject matter. Students were given five minutes to complete each questionnaire, with the exception of the fourth graders, who were given more time because the items in the general questionnaire were read aloud for them. The third questionnaire, newly developed for 1992,

Table 4-13
Main Sample Booklet Contents and Number of Booklets Administered
Age 9/grade 4, Writing

Subject Area	Booklet Number	Common Background Block	Cognitive Blocks		Subject Area Background Block	Motivation Background Block	Number of Booklets Administered
Writing	60	CA	W6	W7	W2	WX	536
	61	CA	W7	W8	W2	WX	530
	62	CA	W8	W6	W2	WX	538
	63	CA	W3	W4	W2	WX	524
	64	CA	W4	W5	W2	WX	528
	65	CA	W5	W3	W2	WX	532
	66	CA	W9	W10	W2	WX	531
	67	CA	W10	W11	W2	WX	530
	68	CA	W11	W9	W2	WX	539
	69	CA	W6	W4	W2	WX	539
	70	CA	W5	W7	W2	WX	525
	71	CA	W8	W3	W2	WX	528
	72	CA	W3	W10	W2	WX	527
	73	CA	W4	W11	W2	WX	531
	74	CA	W9	W5	W2	WX	537
	75	CA	W7	W9	W2	WX	526
	76	CA	W10	W8	W2	WX	535
	77	CA	W11	W6	W2	WX	516

Table 4-14
Cognitive and Noncognitive Block Information
Age 9/grade 4, Writing

Block	Type	Total Number of Items	Number of Multiple-choice Items	Number of Extended Constructed-response Items	Booklets Containing Block
CA	Common Background	20	20	—	60 - 77
WB	Writing Background	16	16	—	60 - 77
WX	Motivation Background	5	5	—	60 - 77
W3	Writing Cognitive (Informative)	1	0	1	63, 65, 71, 72
W4	Writing Cognitive (Informative)	1	0	1	63, 64, 69, 73
W5	Writing Cognitive (Informative)	1	0	1	64, 65, 70, 74
W6	Writing Cognitive (Narrative)	1	0	1	60, 62, 69, 77
W7	Writing Cognitive (Narrative)	1	0	1	60, 61, 70, 75
W8	Writing Cognitive (Narrative)	1	0	1	61, 62, 71, 76
W9	Writing Cognitive (Persuasive)	1	0	1	66, 68, 74, 75
W10	Writing Cognitive (Persuasive)	1	0	1	66, 67, 72, 76
W11	Writing Cognitive (Persuasive)	1	0	1	67, 68, 73, 77

Table 4-15
Main Sample Booklet Contents and Number of Booklets Administered
Age 13/grade 8, Writing

Subject Area	Booklet Number	Common Background Block	Cognitive Blocks*		Subject Area Background Block	Motivation Background Block	Number of Booklets Administered
Writing	60	CA	W6	W7	WB	WX	561
	61	CA	W7	W8	WB	WX	560
	62	CA	W8	W6	WB	WX	563
	63	CA	W3	W4	WB	WX	580
	64	CA	W4	W5	WB	WX	586
	65	CA	W5	W3	WB	WX	581
	66	CA	W9	W10	WB	WX	560
	67	CA	W10	W11	WB	WX	564
	68	CA	W11	W9	WB	WX	584
	69	CA	W6	W4	WB	WX	582
	70	CA	W5	W7	WB	WX	582
	71	CA	W8	W3	WB	WX	576
	72	CA	W3	W10	WB	WX	595
	73	CA	W4	W11	WB	WX	582
	74	CA	W9	W5	WB	WX	572
	75	CA	W7	W9	WB	WX	571
	76	CA	W10	W8	WB	WX	564
	77	CA	W11	W6	WB	WX	563
	78	CA	W12		WB	WX	2294
	79	CA	W13		WB	WX	2322

---

* Blocks W12 and W13 are 50-minute blocks.

Table 4-16
Cognitive and Noncognitive Block Information
Age 13/grade 8, Writing

Block	Type	Total Number of Items	Number of Multiple-choice Items	Extended Constructed-response Items	Booklets Containing Block
CA	Common Background	22	22	—	60 - 79
WB	Writing Background	26	26	—	60 - 79
WX	Motivation Background	5	5	—	60 - 79
W3	Writing Cognitive (Informative)	1	0	1	63, 65, 71, 72
W4	Writing Cognitive (Informative)	1	0	1	63, 64, 69, 73
W5	Writing Cognitive (Informative)	1	0	1	64, 65, 70, 74
W6	Writing Cognitive (Narrative)	1	0	1	60, 62, 69, 77
W7	Writing Cognitive (Narrative)	1	0	1	60, 61, 70, 75
W8	Writing Cognitive (Narrative)	1	0	1	61, 62, 71, 76
W9	Writing Cognitive (Persuasive)	1	0	1	66, 68, 74, 75
W10	Writing Cognitive (Persuasive)	1	0	1	66, 67, 72, 76
W11	Writing Cognitive (Persuasive)	1	0	1	67, 68, 73, 77
W12	Writing Cognitive (50-minute Informative)	1	0	1	78
W13	Writing Cognitive (50-minute Narrative)	1	0	1	79

Table 4-17
Main Sample Booklet Contents and Number of Booklets Administered
Age 17/grade 12, Writing

Subject Area	Booklet Number	Common Background Block	Cognitive Blocks*		Subject Area Background Block	Motivation Background Block	Number of Booklets Administered
Writing	60	CA	W6	W7	WB	WX	539
	61	CA	W7	W8	WB	WX	530
	62	CA	W8	W6	WB	WX	514
	63	CA	W3	W4	WB	WX	537
	64	CA	W4	W5	WB	WX	523
	65	CA	W5	W3	WB	WX	517
	66	CA	W9	W10	WB	WX	513
	67	CA	W10	W11	WB	WX	517
	68	CA	W11	W9	WB	WX	523
	69	CA	W6	W4	WB	WX	532
	70	CA	W5	W7	WB	WX	524
	71	CA	W8	W3	WB	WX	540
	72	CA	W3	W10	WB	WX	518
	73	CA	W4	W11	WB	WX	519
	74	CA	W9	W5	WB	WX	538
	75	CA	W7	W9	WB	WX	535
	76	CA	W10	W8	WB	WX	522
	77	CA	W11	W6	WB	WX	526
	78	CA	W12		WB	WX	2038
	79	CA	W13		WB	WX	2079
	80	CA	W14		WB	WX	2085

---

* Blocks W12, W13, and W14 are 50-minute blocks.

Table 4-18
Cognitive and Noncognitive Block Information
Age 17/grade 12, Writing

Block	Type	Total Number of Items	Number of Multiple-choice Items	Number of Extended Constructed-response Items	Booklets Containing Block
CA	Common Background	29	29	—	60 - 80
WB	Writing Background	26	26	—	60 - 80
WX	Motivation Background	5	5	—	60 - 80
W3	Writing Cognitive (Informative)	1	0	1	63, 65, 71, 72
W4	Writing Cognitive (Informative)	1	0	1	63, 64, 69, 73
W5	Writing Cognitive (Informative)	1	0	1	64, 65, 70, 74
W6	Writing Cognitive (Narrative)	1	0	1	60, 62, 69, 77
W7	Writing Cognitive (Narrative)	1	0	1	60, 61, 70, 75
W8	Writing Cognitive (Narrative)	1	0	1	61, 62, 71, 76
W9	Writing Cognitive (Persuasive)	1	0	1	66, 68, 74, 75
W10	Writing Cognitive (Persuasive)	1	0	1	66, 67, 72, 76
W11	Writing Cognitive (Persuasive)	1	0	1	67, 68, 73, 77
W12	Writing Cognitive (50-min. Info.)	1	0	1	78
W13	Writing Cognitive (50-min. Narr.)	1	0	1	79
W14	Writing Cognitive (50-min. Pers.)	1	0	1	80

Table 4-19
Long-term Trend Sample Booklet Contents and Number of Booklets Administered
Age Class 9

Subject Area	Booklet Number	Common Background Block	Subject Area Background Block	Cognitive Blocks			Number of Booklets Administered
Reading and Writing	51	BZ	*	BC	BL	BQ	1186
	52	BZ	*	BH	BE	BR	1165
	53	BZ	*	BC	BK	BJ	1178
	54	BZ	*	BG	BO	BE	1180
	55	BZ	*	BM	BG	BN	1169
	56	BZ	*	BV	BR		1184
Mathematics and Science	91	C1	*	R1	M1	S1	2388
	92	C1	*	S2	R2	M3**	2512
	93	C1	*	M2	S3	R3	2435

---

* Subject area background questions are included in cognitive blocks for this booklet.

** Calculator needed for this block.

Table 4-20
Long-term Trend Sample Booklet Contents and Number of Booklets Administered
Age Class 13

Subject Area	Booklet Number	Common Background Block	Subject Area Background Block	Cognitive Blocks			Number of Booklets Administered
Reading and Writing	51	BZ	*	BM	BK	BD	919
	52	BZ	*	BC	BL	BQ	906
	53	BZ	*	BH	BE	BR	923
	54	BZ	*	BN	BC	BD	905
	55	BZ	*	BG	BO	BE	928
	56	BZ	*	BG	BJ	BP	933
Mathematics and Science	91	C1	*	R1	M1	S1	1928
	92	C1	*	S2	R2	M3**	1976
	93	C1	*	M2	S3	R3	2005

---

* Subject area background questions are included in cognitive blocks for this booklet.

** Calculator needed for this block.

Table 4-21
Long-term Trend Sample Booklet Contents and Number of Booklets Administered
Age Class 17

Subject Area	Booklet Number	Common Background Block	Subject Area Background Block	Cognitive Blocks			Number of Booklets Administered
Reading and Writing	51	BZ	*	BM	BK	BD	927
	52	BZ	*	BC	BL	BQ	924
	53	BZ	*	BH	BE	BR	917
	54	BZ	*	BN	BC	BD	951
	55	BZ	*	BG	BO	BE	939
	56	BZ	*	BG	BJ	BP	911
Mathematics and Science	84	C1	*	M1	M2	S3	2207
	85	C1	*	S1	S2	M3**	2152

---

* Subject area background questions are included in cognitive blocks for this booklet.

** Calculator needed for this block.

Table 4-22
Long-term Trend Sample Block Information, Age Class 9

Block	Type	Total Number of Items	Number of Cognitive Items	Number of Open-ended Items		Booklets Containing Block
				Cognitive	Noncognitive	
BZ	Common Background	37	0	0	1	51 - 56
C1	Common Background	28	0	0	0	91 - 93
BC	Writing Background/Cognitive	23	1	1	0	51, 53
BE	Writing Background/Cognitive	11	2	2	0	52, 54
BG	Writing Background/Cognitive	8	2	2	0	54, 55
BH	Reading Background/Cognitive	15	11	1	0	52
BJ	Reading Background/Cognitive	24	13	1	0	53
BK	Reading Background/Cognitive	19	11	0	0	53
BL	Reading Background/Cognitive	26	7	1	1	51
BM	Reading Background/Cognitive	16	12	1	0	55
BN	Reading Background/Cognitive	25	14	1	0	55
BO	Reading Background/Cognitive	22	11	0	0	54
BQ	Reading Background/Cognitive	21	12	0	0	51
BR	Reading Background/Cognitive	16	12	0	0	52, 56
BV	Reading and Writing Background/Cognitive	36	7 Rd. 1 Wr.	1 Rd. 1 Wr.	0	56
R1	Reading Background/Cognitive	20	9	0	0	91
R2	Reading Background/Cognitive	20	11	0	0	92
R3	Reading Background/Cognitive	17	10	1	0	93
M1	Mathematics Background/Cognitive	26	26	9	0	91
M2	Mathematics Background/Cognitive	26	26	9	0	93
M3	Mathematics Background/Cognitive (Calc.)	19	16	10	0	92
S1	Science Background/Cognitive	23	18	0	0	91
S2	Science Background/Cognitive	25	25	0	0	92
S3	Science Background/Cognitive	31	20	0	0	93

BEST COPY AVAILABLE

Table 4-23

Long-term Trend Sample Block Information, Age Class 13

Block	Type	Total Number of Items	Number of Cognitive Items	Number of Open-ended Items		Booklets Containing Block
				Cognitive	Noncognitive	
BZ	Common Background	37	0	0	1	51 - 56
C1	Common Background	30	0	0	0	91 - 93
BC	Writing Background/Cognitive	23	1	1	0	52, 54
BD	Writing Background/Cognitive	25	1	1	0	51, 54
BE	Writing Background/Cognitive	11	2	2	0	53, 55
BG	Writing Background/Cognitive	8	2	2	0	55, 56
BH	Reading Background/Cognitive	18	13	1	1	53
BJ	Reading Background/Cognitive	24	14	2	0	56
BK	Reading Background/Cognitive	17	9	1	0	51
BL	Reading Background/Cognitive	27	6	1	1	52
BM	Reading Background/Cognitive	16	12	1	0	51
BN	Reading Background/Cognitive	23	12	1	0	54
BO	Reading Background/Cognitive	21	10	2	0	55
BP	Reading Background/Cognitive	15	9	1	0	55
BQ	Reading Background/Cognitive	23	17	0	0	52
BR	Reading Background/Cognitive	19	15	0	0	53
R1	Reading Background/Cognitive	31	12	1	0	91
R2	Reading Background/Cognitive	19	10	0	0	92
R3	Reading Background/Cognitive	28	13	0	0	93
M1	Mathematics Background/Cognitive	51	37	9	0	91
M2	Mathematics Background/Cognitive	44	37	8	0	93
M3	Mathematics Background/Cognitive (Calc.)	32	24	10	0	92
S1	Science Background/Cognitive	36	25	0	0	91
S2	Science Background/Cognitive	40	27	0	0	92
S3	Science Background/Cognitive	36	27	0	0	93

Table 4-24

Long-term Trend Sample Block Information, Age Class 17

Block	Type	Total Number of Items	Number of Cognitive Items	Number of Open-ended Items		Booklets Containing Block
				Cognitive	Noncognitive	
BZ	Common Background	48	0	0	1	51 - 56
C1	Common Background	48	0	0	0	84, 85
BC	Writing Background/Cognitive	23	1	1	0	52, 54
BD	Writing Background/Cognitive	25	1	1	0	51, 54
BE	Writing Background/Cognitive	11	2	2	0	53, 55
BG	Writing Background/Cognitive	8	2	2	0	55, 56
BH	Reading Background/Cognitive	19	13	1	2	53
BJ	Reading Background/Cognitive	17	6	2	1	56
BK	Reading Background/Cognitive	17	9	1	0	51
BL	Reading Background/Cognitive	32	6	1	2	52
BM	Reading Background/Cognitive	36	12	1	0	51
BN	Reading Background/Cognitive	32	12	1	1	54
BO	Reading Background/Cognitive	24	13	1	0	55
BP	Reading Background/Cognitive	25	11	1	0	56
BQ	Reading Background/Cognitive	17	11	1	0	52
BR	Reading Background/Cognitive	20	9	0	0	53
M1	Mathematics Background/Cognitive	49	35	10	0	84
M2	Mathematics Background/Cognitive	49	35	5	0	84
M3	Mathematics Background/Cognitive (Calc.)	35	24	14	0	85
S1	Science Background/Cognitive	38	27	0	0	85
S2	Science Background/Cognitive	41	32	0	0	85
S3	Science Background/Cognitive	32	23	0	0	84

BEST COPY AVAILABLE

followed the three cognitive blocks and contained five questions about students' motivation to do well on the assessment, their perceptions concerning the difficulty of the assessment, and their familiarity with types of questions included.

4.3.1 Reading Background Questionnaires

Three categories of information were represented in the five-minute sections of reading background questions called the **student reading questionnaire** (14 questions at grade 4, 24 questions at grades 8 and 12):

> *Time Spent Studying Reading*: Time spent on task and reading coursework has been shown to be strongly related to reading achievement (Anderson, Hiebert, Scott, & Wilkinson, 1984). Students were asked to describe both the amount of instruction they received in reading and the time spent on reading homework.
>
> *Instructional Practices*: The nature of students' reading instruction is also thought to be related to achievement (Dole, Duffy, Roehler, & Pearson, 1991). Students were asked to report their instructional experiences related to reading in the classroom, including group work, special projects, and writing in response to reading. In addition, they were asked about the instructional practices of their reading teachers and the extent to which the students themselves discussed what they read in class and demonstrated use of skills and strategies.
>
> *Attitudes Towards Reading*: Students' enjoyment of and confidence in their abilities in reading and their perceptions of the usefulness of reading to their present and future lives appear to be related to reading achievement (Guthrie & Greaney, 1991). Students were asked a series of questions about their attitudes and perceptions about reading, such as whether they enjoyed reading and whether they were good in reading.

4.3.2 Mathematics Background Questionnaire

Three categories of information were represented in the five-minute **student mathematics questionnaire** (18 questions at grade 4 and 23 questions at grades 8 and 12): time spent on task and mathematics coursework, the nature of students' mathematics instruction, and students' enjoyment of and confidence in their abilities in mathematics and their perceptions of the usefulness of the discipline to their present and future lives. This questionnaire was the second section in every student assessment booklet.

4.3.3 Writing Background Questionnaire

Students participating in the writing assessment answered a set of general background questions and a set of writing-specific questions. Together, these sets of questions were designed to gather contextual information about students, the nature of their writing instruction and writing practices, and their attitudes toward writing. A one-minute questionnaire was given to students at the end of each booklet to determine their motivation in completing the assessment and their familiarity with assessment tasks. In order to ensure that all grade 4 students understood the questions and had every opportunity to answer them, the three sets of questionnaires were read aloud by administrators as students read along and responded in their booklets.

4.4 TEACHER QUESTIONNAIRES

4.4.1 Reading Teacher Questionnaire

To supplement the information on instruction reported by students, the reading teachers of the fourth graders participating in the NAEP reading assessment were asked to complete a questionnaire about their instructional practices, teaching backgrounds, and characteristics. The teacher questionnaire contained two parts. The first part pertained to the teachers' background and general training. The second part pertained to specific training in teaching reading and the procedures the teacher used for *each class* containing an assessed student.

The Teacher Questionnaire, Part I: Background and General Training (23 questions) included questions pertaining to gender, race/ethnicity, years of teaching experience, certification, degrees, major and minor fields of study, coursework in education, coursework in specific subject areas, amount of in-service training, extent of control over instructional issues, and availability of resources for their classroom. This component of the questionnaire was completed by teachers whose students participated in any subject assessed in NAEP.

The **Teacher Questionnaire, Part II: Training in Reading and Classroom Instructional Information** (56 questions) included questions on the teacher's exposure to various issues related to reading and teaching reading through pre- and in-service training, ability level of students in the class, whether students were assigned to the class by ability level, time on task, homework assignments, frequency of instructional activities used in class, methods of assessing student progress in reading, instructional emphasis given to the reading abilities covered in the assessment, and use of particular resources. This section of the questionnaire was completed only by teachers whose students took part in the reading assessment.

4.4.2 Mathematics Teacher Questionnaire

To supplement the information on instruction reported by students, the mathematics teachers of the fourth- and eighth-grade students participating in the assessment were asked to complete a mathematics teacher questionnaire about their instructional practices, teaching backgrounds, and characteristics. The teacher questionnaires contained two parts.

The Teacher Questionnaire, Part I: Background and Training (23 questions at grade 4 and 32 questions at grade 8) included questions pertaining to gender, race/ethnicity, years of teaching experience, certification, degrees, major and minor fields study, coursework in education, coursework in subject area, in-service training, extent of control over classroom, instruction, and curriculum, and availability of resources for their classroom.

The **Teacher Questionnaire, Part II: Class by Class Mathematics Information** (40 questions at grade 4 and 42 questions at grade 8) pertained to the procedures the teacher used for *each class* containing an assessed student and included questions on the ability level of students in the class, whether students were assigned to the class by ability level, time on task, homework assignments, frequency of instructional activities used in class, instructional emphasis given to the topics and skills covered in the assessment, and use of particular resources.

4.4.3 Writing Teacher Questionnaire

A teacher questionnaire was administered to the English or language arts teachers of eighth graders participating in the assessment. This questionnaire contained two parts. The first requested information about the teacher, such as race/ethnicity and gender, as well as academic degrees held, teaching certification, training in writing, and ability to get instructional resources. In the second part, teachers were asked to provide information on each class they taught that included one or more students who participated in the assessment. The information included, among other things, the amount of time spent on writing instruction and homework, the extent to which various writing assignments were given, the instructional and grading emphases placed on different aspects of writing, and the use of various instructional approaches (e.g., peer response and computers).

4.5 EXCLUDED STUDENT QUESTIONNAIRE

Some students selected for the assessment were judged by school authorities to be incapable of participating in the assessment because they had limited English language proficiency, were mildly mentally retarded (educable), or were functionally disabled.

For each student excluded from the assessment, schools were required to complete a questionnaire containing 27 questions about the characteristics of that student and the reason for exclusion. For students with an Individual Education Plan, the questionnaire included questions about students' functional grade level, mainstreaming, and special education programs. For Limited English Proficient students, it asked about students' native language, time spent in special education and language programs, and the level of the students' English language proficiency.

4.6 SCHOOL CHARACTERISTICS AND POLICIES QUESTIONNAIRE

A school characteristics and policy questionnaire was given to the principal or other administrator of each school that participated in the 1992 NAEP assessment. This questionnaire asked questions (77 at age classes 9 and 13, 76 at age class 17) about background and characteristics of school principals, length of school day and year, school enrollment, absenteeism, drop-out rates, size and composition of teaching staff, policies about tracking, curriculum, testing practices and use, special priorities and school-wide programs, availability of resources, special services, community services, policies for parental involvement, and school-wide problems.

4.7 PRINCIPAL QUESTIONNAIRE

Before the assessment, Westat, Inc., distributed a questionnaire to the principal of each participating school to gather data about school characteristics, including school enrollment and attendance, parents' occupations, and student race/ethnicity. These data were used in part to estimate the number of age/grade-eligible students and to determine the correct "size and type of community" classification for each school.

4.8 THE WRITING PORTFOLIO STUDY

As part of the 1992 writing trend assessment, NAEP conducted a special study of school-based writing, involving random samples of subsets of fourth- and eighth-grade students who were assessed in writing. The purpose of this "portfolio" study was to expand the view of students' writing abilities by analyzing pieces of writing that they produced outside of the assessment situation. The primary writing teachers of sampled students were asked to assist students in selecting three pieces of writing that represented their best writing efforts, a range of writing tasks, and the use of writing process strategies. In addition, students were asked to write a letter to NAEP explaining why they chose their three pieces and what they like about each piece.

Chapter 5

FIELD OPERATIONS AND DATA COLLECTION

Nancy W. Caldwell and Lesly Flemming

Westat, Inc.

As a subcontractor to Educational Testing Service on the National Assessment of Educational Progress (NAEP), Westat, Inc. was responsible for selecting schools and student samples, field operations, and data collection for the 1992 national component of NAEP. This chapter provides an overview of the field and data collection activities. A detailed discussion of these activities is given in the *Report on Field Operations and Data Collection Activities, 1992 National Assessment of Educational Progress* (Caldwell & Flemming, 1993).

In 1992, the national assessment comprised main samples and long-term trend samples. In general, the main assessments involved new items and components; in the long-term trend assessments, elements of previous years' assessments were carried forward to provide data for reporting on trends in achievement. A total of 773 public and private schools took part in the long-term trend assessments of reading, mathematics, science, and writing, which were administered between October 7, 1991, and May 15, 1992. A total of 1,582 public and private schools took part in the main assessments of reading, mathematics, and writing, which were administered between January 6 and April 3, 1992. (Some schools took part in both main and long-term trend assessments.) Aspects of the field administration and data collection activities for each component are described in the sections that follow.

5.1 SCHEDULE OF ACTIVITIES

Recruiting of schools for NAEP began in May 1991, once the sample of schools had been selected and their corresponding school districts identified. The National Center for Education Statistics and Educational Testing Service contacted the Chief State School Officers in each state notifying them of the districts and schools in their states which were in the sample. In the 35 of the 44 jurisdictions participating in the Trial State Assessment that also had schools sampled for the National Assessment, the state coordinator was sent the list of districts and schools sampled for both the National and the Trial State Assessments.

Throughout June and July 1991, ETS sent NAEP materials to superintendents and heads of private schools inviting their participation. It was during this period that districts were sent a list of their sampled schools. These initial contacts, which were completed prior to supervisor training, paved the way for the telephone contacts by NAEP field supervisors to follow.

Summarized in Table 5-1 are the scheduled project activities for the 1992 National Assessment.

Table 5-1
Summary of Field Operations and Data Collection Schedule

Date	Activity
Mid-May 1991	Department of Education sent first letter to Chief State School Officers about the 1992 assessment. Westat sent state coordinators a list of their schools initially selected for National and/or Trial State Assessments.
June-July 1991	ETS sent NAEP materials to districts to begin the process of obtaining the cooperation of selected schools.
August 21-24, 1991	Training session was held for long-term trend assessment supervisors and main assessment schedulers.
Early September to Mid-December 1991	Supervisors contacted district superintendents and private school principals about the National Assessment, to establish or confirm participation and to make arrangements for introductory meetings with school representatives. Westat sent letter to superintendents confirming the date, time, and place of introductory meeting. Package of materials was sent to principals of participating schools. Supervisors conducted introductory meetings for the assessment. Westat selected substitutes for refusing schools.
October 7 - December 13, 1991	Fall long-term trend assessments were conducted.
December 4 - 7, 1991	Main assessment supervisor training session was held.
January 6 - March 13, 1992	Winter long-term trend assessments were conducted.
January 6 - April 3, 1992	Main assessments were conducted.
March 16 - May 15, 1992	Spring long-term trend assessments were conducted.

5.2 TRAINING OF FIELD STAFF

5.2.1 Supervisor Training

Training for National Assessment supervisors was multiphased and involved separate sessions conducted in August and December 1991. All training was conducted by the Westat project director and field director, with assistance from home office field managers. Also in attendance were representatives from ETS, National Computer Systems (NCS), and the

National Center for Education Statistics. The first of these training sessions was held August 21-24, 1991, in Arlington, Virginia. Attending were the 14 supervisors responsible for conducting trend assessments, 30 NAEP-experienced supervisors who would work throughout the fall period to gain cooperation and schedule assessment activities with main NAEP schools, five "troubleshooters," and the field managers.

Topics discussed at the first training session included an overview of NAEP; supervisory responsibilities; a summary of various reports from recent assessments; procedures for contacting districts and conducting introductory meetings, scheduling assessments within the PSUs, and recruiting and training exercise administrators; classroom management techniques; and administrative forms and procedures. In addition, long-term trend supervisors and troubleshooters received specific instruction on procedures for drawing the sample of students, conducting assessments, and preparing and distributing assessment questionnaires. Also featured were practice exercises in sampling and in filling out the various administrative forms. A mock assessment session was held with the supervisors acting the role of students. This included reading verbatim from one of the actual assessment scripts and following the prescribed procedures for distributing materials, reading directions, and recording the results of the assessment.

The 57 NAEP supervisors who would be responsible for main NAEP assessment activities were trained during the second session, held December 4-8, 1991. Training focused on a review of the preliminary activities during the fall (results of initial contacts with districts and schools, introductory meetings, scheduling of assessments, and the status of exercise administrator recruitment), and assessment activities (sampling, excluding students, teacher surveys, and administrative forms and procedures). In response to a request from supervisors to address issues on classroom management, the December training session included a presentation by a guest speaker from the Howard County, Maryland, Substitute Teacher Center.

5.2.2 Recruiting, Hiring, and Training Exercise Administrators

During the fall, while the supervisors were securing cooperation in districts and schools and scheduling assessments, they also recruited and hired exercise administrators. The exercise administrator's primary job was to administer the assessment sessions. Exercise administrators were recruited from many sources. Each supervisor was given a PSU-by-PSU computer list of interviewers and exercise administrators who had worked for Westat on other studies, including NAEP. During their initial contacts with schools, the supervisors asked the school principals and other staff to recommend potential exercise administrators. Where necessary, advertisements were placed in local newspapers and an employment service was notified.

Supervisors were encouraged to hire locally and to hire individuals with teaching experience or the ability to handle classroom situations. Many of the 630 exercise administrators hired in 1992 were retired or substitute teachers.

The assessment supervisors had complete responsibility for recruiting, hiring, training, and supervising their exercise administrators. The Supervisor Manual discussed the training and use of exercise administrators in conducting assessments. In addition, one session of the supervisors' training included a discussion of exercise administrator training and a thorough review of the Exercise Administrator Manual. The supervisors gave a copy of the manual to each exercise administrator before the training session was held.

Exercise administrators were required to study the manual first and attend a half-day training session conducted by the supervisor. During the training, the supervisor reviewed, in detail, all aspects of the administrator's job, including preparing materials, booklets, and Administration Schedules for assessment; the actual conduct of the session; post-assessment collection of booklets, pencils, and other assessment materials; coding booklet covers; recordkeeping; and administrative matters.

5.3 SECURING COOPERATION IN SAMPLED SCHOOLS

After the August training session, supervisors began working on obtaining cooperation. The approach the supervisors took when calling superintendents depended on whether the district had been notified about NAEP by the state coordinator and whether the district also had schools selected for the Trial State Assessment. For districts that had been contacted by the state coordinator, the supervisor began by referring to that contact. If the district also had schools sampled for the Trial State Assessment, the supervisor explained that the introductory meeting was to describe the schools' role in the National Assessment and that only representatives from those schools sampled for the National Assessment should attend.

As the supervisors contacted superintendents and private school officials to establish cooperation and to set up the introductory meeting, they notified Westat's home office. As introductory meetings were scheduled, Westat mailed confirmation letters and packages of informational materials to district offices and school personnel.

5.3.1 Introductory Meetings

From early September through the beginning of December 1991, supervisors visited all 94 PSUs and conducted introductory meetings. The supervisors performed a number of tasks during the introductory meetings. During the meetings supervisors tailored the information given to the audience since school personnel representing different age/grade levels were often present, as well as personnel representing schools in which different forms of the assessment would be conducted. While the content of the meetings varied, they generally included the following:

- presenting an overview of NAEP, using a scripted slide presentation;
- answering questions;
- explaining the tasks that were required of each school;
- setting preliminary sampling and assessment dates for each school;
- collecting and checking completed Principal Questionnaires;

- verifying information about the school and completing the School Control Form;
- distributing appropriate Student Listing Forms and explaining the method of completion;
- identifying a school coordinator (if not already identified); and
- inquiring about possible exercise administrator candidates, where needed.

In general, introductory meetings lasted about one hour. They ranged in size from small meetings between the supervisor and one school coordinator to formal meetings attended by 20 to 30 school officials (superintendents, curriculum specialists, testing personnel, principals, and coordinators). The introductory meetings were the first opportunity for principals and other officials at the school level to discuss the National Assessment with NAEP staff. Thus, the meetings were particularly important for establishing rapport with the schools, assuring school cooperation, and explaining the details of the schools' tasks to the individuals responsible for them.

The usual procedure was for the supervisor to conduct the meetings alone. However, for those meetings where a large number of district and school representatives were expected, or for those schools and districts that were considered to be problematic, ETS staff were requested to take the lead in the introductory meetings.

5.3.2 Making Arrangements for the Assessments

During the introductory meetings, the supervisor discussed arrangements for the assessments with representatives from each school. Within the weeks scheduled for each PSU, the supervisor had the flexibility to set each school's assessment date in coordination with school staff. The staff sometimes expressed preferences for a particular day or dates or had particular times when the assessment could not be scheduled. Their preferences or restrictions depended on the events that had already been scheduled on their school calendar. Using this information from the schools, the supervisors set up the assessment schedule for the PSU.

The School Control Form was used by the supervisors to record information about the school's assessment plan. The form gave estimates of the number of students to be assessed in the school as well as the type of sessions (print-administered and/or tape-administered) to be held. Using this information, the supervisor and school staff could discuss the approximate number of sessions to be held in the school and the space required.

The supervisor usually learned during the introductory meeting whether a school required some form of parental notification or permission. In preparation for this, the supervisor had copies of three versions of standard NAEP letters to parents. These letters were made available to schools requesting them. The first version informed parents about the assessment. The second version assumed parental consent unless parents sent the form back stating they did not want their child to participate in the assessment. The third version required that parents sign and return the form before students could be assessed. Schools were offered their choice of the letters, although when the issue of parental permission came up in discussions, supervisors offered the least restrictive version first. Schools could send out their own letters and notices if they preferred not to use the ones prepared by NAEP.

Following an introductory meeting, the supervisor sent the Principal Questionnaires and copies of the School Control Forms to the home office.

5.4 LONG-TERM TREND ASSESSMENTS

5.4.1 Overview

To provide continuity and comparability with past NAEP studies, the long-term trend assessments replicated procedures and materials that were used in NAEP assessments prior to 1992. Student eligibility was also based on criteria used in years prior to 1988 when the modal grade changed from the eleventh grade to the twelfth. Thus, the spring long-term trend assessment, held in the nine-week period between March 16 and May 15, 1992, included age 17/grade 11 students. The fall assessment of age 13/grade 8 students was held in the ten-week period between October 7 and December 13, 1991 and the winter assessment of age 9/grade 4 students was held during the ten-week period from January 6 to March 13, 1992.

Tape-administered sessions were conducted with samples of age-eligible students, as had been done in all previous years. Additional samples of age- and grade-eligible students were assessed with print-administered booklets, following procedures initiated in the 1984 assessment. Six different types of sessions were used in the long-term trend assessments—one print-administered and five different tape-administered sessions. Depending on the size of a participating school, up to four different session types might be conducted there.

Table 5-2 gives summary information for the types of sessions conducted for the long-term trend assessments.

Table 5-2
Session Summary, Long-term Trend Assessments

Age/Grade	Session Type	Booklets Used	Subjects	Materials Used
Age 9/Grade 4 (Winter)	Print Tape	51 - 56 91 - 93	Reading, Writing Reading, Mathematics, Science	— Calculators (Booklet 92)
Age 13/Grade 8 (Fall)	Print Tape	51 - 56 91 - 93	Reading, Writing Reading, Mathematics, Science	— Calculators (Booklet 92)
Age 17/Grade 11 (Spring)	Print Tape	51 - 56 84, 85	Reading, Writing Reading, Mathematics, Science	— Calculators (Booklet 85)

5.4.2 Selecting the Student Sample

Two weeks prior to a school's assessment date, the NAEP supervisor contacted the school coordinator to make sure that the list of eligible students was prepared and that all agreed arrangements were set. The supervisor visited the school (or district office) before the assessment date to select the sample of students. The time interval between the selection of the sample and assessment varied depending on several factors, most notably the size of the school. The average elapsed time was about a week.

The supervisor's first task upon arriving at the school to select the sample was to review the Student Listing Form (or comparable list of students) to be sure that they had been completed correctly. The supervisor made certain checks to assure that all age- and grade-eligible students had been listed. The supervisor also checked that the students to be excluded from the assessment were listed so that they could be included in the sample.

For each school, the Westat home office produced a Session Assignment Form, which told the supervisor how to select the sample in that school. The Session Assignment Form contained the following information:

- **Identifying Information** - identifying the school, the age and grade level, and approximate number of students to be assessed in total and by session type.
- **Type of Session** - specifying whether the school was to have only print-administered sessions, only tape-administered sessions, or both. If tape-administered sessions were assigned to the schools, then the particular types of tape-administered sessions were specified.
- **Sampling Instructions** - the particular steps the supervisor should follow in selecting the sample of eligible students in the school:

    1) Review the lists for completeness.

    2) Consecutively number all eligible students.

    3) Compare the number of eligibles to the minimum and maximum specified on the Session Assignment Form. If the actual number of eligibles was out of range, the supervisor called Westat for additional instructions.

    4) Select the sample of students as specified on the Session Assignment Form. The steps to be followed depended on whether the school was selected for print-administered sessions, tape-administered sessions, or both.

If the school was selected for tape-administered sessions (for which only age-eligible students are considered), the supervisor marked off any student who did not meet the age eligibility criteria after marking the selected line numbers. The eligible students were then assigned to the appropriate tape-administered sessions.

Following the sampling instructions, the supervisor was instructed to fill out Administration Schedules, listing the sampled students for each session. Before completing the list, the supervisor reviewed the plans for the assessment with the school coordinator. If, for example, a large number of students was sampled for a print-administered session, the supervisor discussed NAEP's preference for this group to be divided into sessions of about 30 students each. Also discussed were procedures which might be helpful to the school, such as listing students on the Administration Schedules alphabetically or by homeroom. Sometimes the coordinator had very specific ideas about the organization of the assessment.

After the excluded students were identified, the supervisors were instructed to prepare and distribute the Excluded Student Questionnaires. If the coordinator could not identify the excluded students while the supervisor was at the school, a set of Instructions For Excluding Students (Figure 5-1) was left with the coordinator along with an estimated number of questionnaires needed.

The supervisor and exercise administrators were encouraged to assign booklets to the sampled students before leaving the school, recording the booklet numbers on the Administration Schedules. Then they clearly labeled the bundle of preassigned booklets for each session so that they could be removed from the school and matched with the appropriate Administration Schedule when the supervisor and exercise administrators returned for the assessment. If requested by the coordinator, the supervisor and administrators would fill out Appointment Cards for each student telling him or her when and where to report for the assessment and prepare lists of sampled students for distribution to teachers. They would also prepare parental notification or consent letters if requested by the school.

5.4.3 Conduct of the Assessment

The primary responsibility for conducting assessment sessions was with the exercise administrators. Supervisors were required to observe the first session each exercise administrator conducted to ensure that he or she followed the procedures properly. Supervisors were also required to be present in all schools during the assessments, if at all possible, especially in large schools with several sessions. Previous experience has shown that the supervisor plays an important role as the liaison between the National Assessment and school staff ensuring that the assessments go smoothly. If the supervisor is present, he or she can, for example, help direct students to correct rooms when more than one session is being conducted at the same time.

To ensure that sessions were administered in a uniform way, the exercise administrator was provided with scripts for each session type, to be read verbatim. The scripts began with a brief introduction to the study. The exercise administrator was then directed to distribute the booklets, being careful to match each student with his or her preassigned booklet.

Following the distribution of booklets, the scripts differed depending on whether the session was print- or tape-administered. In print-administered sessions, the exercise administrator's script specified the timing of the sections of the booklets. In tape sessions, the exercise administrator was instructed to turn on the tape recorder after distributing the booklets; sections were then administered and timed by the audiotaped script.

Figure 5-1

Instructions for Excluding Students (Long-term Trend Assessments)

These instructions accompany the NAEP Administration Schedules listing the students selected to participate in the National Assessment. Please review the Administration Schedule and, in concert with other school officials, determine whether any students should be excluded from the assessment because they are non-English speaking, educable mentally retarded, or functionally disabled. These categories are defined as follows:

- **Non-English speaking students** - Those who do not read or speak English and would be unable to overcome the language barrier in the test situation.
- **Educable mentally retarded (EMR)** - Students who have been psychologically tested as EMR students or students who are considered EMR in the professional opinion of the principal or other qualified staff members. However, students should not be excluded because of poor academic performance or normal discipline problems. Only those students should be excluded who cannot give meaningful responses to exercises at their age level.
- **Functionally disabled** - (temporary or permanent physical disability) Students who are so disabled that they cannot perform in the NAEP testing situation should also be excluded. However, functionally disabled students who can respond should be included.

If the school determines that a student should be excluded, the student's name should be lined through, taking care to make sure that the name is still legible.

Next, prepare an Excluded Student Questionnaire for each sampled student the school has determined should be excluded from the assessment. Write the student's name in the area labeled "Teacher of -" and give the questionnaire to the staff member most knowledgeable about the student.

The Excluded Student Questionnaire should be returned by the day of the assessment and given to the NAEP Supervisor or Exercise Administrator who will complete the coding on the cover of the questionnaire.

During the sessions, the exercise administrators walked around the room monitoring the students to make sure they were working in the correct section of their booklets and to prevent them from looking at a neighbor's booklet. During the background (first) section, exercise administrators were allowed to assist students in understanding questions and responding to them. After the students began working on the other sections of the booklets, administrators were not allowed to answer any student's question.

At the end of an assessment session, booklets were collected and students dismissed according to the school's policy. The exercise administrator was then responsible for completing the information at the top of the Administration Schedule and coding the covers of all booklets, including those assigned to absent students.

5.4.4 Results of the Long-term Trend Assessments

Of the 1,000 schools sampled for the long-term trend assessments, 764 were assessed. In addition, 11 substitute school were assessed, for a total of 775 assessed schools. The assessed schools included 307 at age 9/grade 4, 251 at age 13/grade 8, and 217 at age 17/grade 11.

A total of 35,748 students participated in the long-term trend assessments: 14,397 at age 9/grade 4; 11,423 at age 13/grade 8; and 9,928 at age 17/grade 11.

The overall student participation rate of 89.6 percent remains high when compared to previous years. Participation rates among age 9/grade 4 students has historically been the highest and remained so in 1992 at 94.0 percent, surpassing the 1990 participation rate of 92.4 percent for this age/grade. At age 13/grade 8, the 1992 student participation rate of 90.8 percent is comparable to the 1990 rate of 90.4 percent. The response rate at age 17/grade 11 of 82.8 percent is an increase over the 81.2 percent assessed in 1990.

A total of 2,525 students (6.3%) were sampled for the long-term trend assessments but excluded from participation by their schools because of limited English-speaking ability or a disability. This exclusion rate is slightly higher than the two previous assessments which had an exclusion rate ranging from 5.5 to 6.0 percent.

5.4.5 Assessment Questionnaires and Reports

The School Characteristics and Policies Questionnaire and the Excluded Student Questionnaire were distributed in the schools to be completed by school personnel.

All schools were mailed the School Characteristics and Policies Questionnaire by Westat prior to the assessment. This form was to be filled out by the principal or other staff member knowledgeable about the school's administrative policies and staff characteristics. The supervisors collected this questionnaire when they were at the school for sampling or for the assessment.

An Excluded Student Questionnaire was to be filled out for every student who was sampled for the assessment but excluded by the school. Following exclusion criteria used in previous assessment years, schools could exclude students with limited English-speaking ability, those who were educable mentally retarded, or functionally disabled students, if in the judgment of school staff or if school records indicated they were unable to take the assessment. After the sample of students was drawn and Administration Schedules prepared, the supervisor requested that the school coordinator identify any students who should be excluded. The supervisor then gave an Excluded Student Questionnaire to the coordinator for every excluded student.

The supervisor attempted to collect all completed school and excluded student questionnaires on the assessment day. If the questionnaires were not ready, and it was convenient for the supervisor or an exercise administrator to return to the school later to pick up the questionnaires, they would do so. Otherwise, the supervisor gave the coordinator a postage-paid envelope to use to mail the forms to NCS. Of the 775 School Characteristics and Policies Questionnaires mailed to schools, 765 were completed and collected by NAEP supervisors. This is a completion rate of 99 percent. The return of Excluded Student Questionnaires was comparably high, at 98 percent: 2,525 questionnaires were distributed and 2,487 of these were returned. These numbers do not include questionnaires mailed to NCS directly by the school coordinators.

Once the assessments were completed in a school, the supervisor and exercise administrators completed the coding of the front covers of the assessment booklets, filled out the necessary forms, and shipped the booklets and forms to NCS. A copy of all forms was sent to Westat so that progress in the field could be closely monitored.

The School Worksheet was used to summarize the results of the assessment sessions in each school. The number of students to be assessed, actually assessed, and absent were entered so that the supervisor could calculate if a makeup session(s) was required. If a makeup was required for one or more session types, the supervisor discussed the scheduling of the makeup with the coordinator. The top (original) copy of the School Worksheet, Roster of Questionnaires, and the Administration Schedules (with the students' names removed and left at the school) were included with the booklets in the shipment to NCS. In addition, the supervisor included the following forms as necessary in each shipment:

> **Packing List** - an NCS form that inventoried assessment materials sent to each school. This list was returned with a school's assessment materials to summarize the total number of used and unused booklets in the shipment.
>
> **NAEP Identification Sheet** - sent with each individual bundle of booklets. This sheet was returned with a school's assessment materials to account for the number of used and unused booklets in each bundle.
>
> **Supplemental Transmittal Form** - used for shipping any questionnaires after the original shipment for the school had been sent.

5.5 MAIN ASSESSMENTS

5.5.1 Overview

The main NAEP assessments were much larger and more complex than the long-term trend assessments. About 117,000 students in 1,582 schools participated in main assessments during the 14-week period from January 3 through April 3, 1992. The three groups of students assessed were age 9/grade 4, age 13/grade 8, and age 17/grade 12. Five different types of sessions were possible—four print-administered (each for a different subject) and one tape-administered. Table 5-3 gives summary information for the types of sessions conducted in the main NAEP assessments.

5.5.2 Selecting the Student Sample

Upon arriving at the school (or district office) to select the sample, the supervisor first reviewed the lists of age- and grade-eligible students. He or she confirmed with the school coordinator that all eligible students were listed. If any eligible students were omitted, sampling could not proceed until problems were corrected.

Using the school's Session Assignment Form, the supervisor selected the sample of students to be assessed. After making sure that all eligibles had been listed, the supervisor numbered the students. If the total number of eligible students was within the minimum and maximum limits indicated on the Session Assignment Form, the supervisor could proceed to select the sample. If the number was outside the limits, he/she called Westat for sampling instructions. The supervisor then proceeded to select the sample of students as specified on the form. The Session Assignment Forms provided step-by-step instructions for sampling, as they did for the long-term trend assessments.

Once students were assigned to sessions, the supervisor and exercise administrators filled out an Administration Schedule for each session. The supervisor discussed the final schedule of the sessions with the coordinator and the date, time, and location of each session were filled in on the Administration Schedules.

The supervisor then asked the coordinator to identify any students with an individualized education plan (IEP) and/or who were designated as limited English proficient (LEP) on the Administration Schedules. Next, the coordinator determined whether any of these students should be excluded from the assessment based on the criteria for excluding students in the main assessment schools (Figure 5-2). For each excluded student, an Excluded Student Questionnaire was prepared. If the school coordinator could not identify the excluded students while the supervisor was at the school, the instructions were left with the coordinator along with some questionnaires.

If requested by the school, the supervisor and/or exercise administrators made lists of the sampled students for the teachers and completed appointment cards notifying students about their assessment schedule. If needed, teacher notification letters were also prepared.

Table 5-3

Session Summary, Main Assessments

Age/Grade	**Subject**	**Session Type**	**Booklets Used**	**Student Materials**	**Administrator Materials**
Age 9/Grade 4	Reading	Print	R30-R45	—	Reading poster
	Writing	Print	W60-W77	—	Writing poster
	Mathematics	Print	M1-M26	Four-function calculators (TI-108), rulers, geometric shapes	Mathematics poster
	Mathematics Calculator Bridge	Print	M28C*	Four-function calculators (Sharp EL-240S)	Calculator poster
	Mathematics	Tape	M27T	—	Tape recorder Tape 27T
Age 13/Grade 8	Reading	Print	R30-R49	NAEP readers, bus schedules	—
	Writing	Print	W60-W79	—	—
	Mathematics	Print	M1-M26	Scientific calculators, protractor-rulers, geometric shapes	Calculator poster, mathematics poster
	Mathematics	Tape	M27T	—	Tape recorder Tape 27T
Age 17/Grade 12	Reading	Print	R30-R50	NAEP readers, bus schedules, tax forms	—
	Writing	Print	W60-W80	—	—
	Mathematics	Print	M1-M26	Scientific calculators, protractor-rulers, geometric shapes	Calculator poster, mathematics poster
	Mathematics	Tape	M27T	—	Tape recorder Tape 27T

---

* The EL-240S calculator was used in trend assessments prior to 1992.

Figure 5-2

Instructions for Excluding Students (Main Assessments)

These instructions accompany the NAEP Administration Schedules listing the students selected to participate in the National Assessment. Please review the Administration Schedules and, in concert with other school officials, identify IEP and LEP students and determine whether any of these students should be excluded from the assessment according to the criteria given below.

First, review the list of sampled students. The Administration Schedule has two columns headed "If IEP (X)" and "If LEP (X)". In the IEP column, enter an "X" for any student who has an Individualized Education Plan (IEP) for reasons other than being categorized as gifted and talented. In the LEP column, enter an "X" for any student who is classified by the school as Limited English Proficient (LEP). Information on IEP and LEP is required for all students regardless of whether they will be excluded from the assessment.

Next identify any students marked IEP or LEP on the Administration Schedule who, in the judgment of appropriate school personnel are unable to take the assessment as follows:

1. **The intent is to assess all selected students. Therefore, all selected students who are capable of participating in the assessment should be assessed.**

2. Some of the students identified on the Administration Schedule as LEP or having an IEP may be incapable of participating meaningfully in the assessment. Staff members knowledgeable about the IEP/LEP students, may exclude such students, as described below.

3. A student identified on the Administration Schedule as LEP may be excluded from the assessment if he/she:

   - Is a native speaker of language other than English; AND
   - Has been enrolled in an English-speaking school (not including a bilingual education program) for less than two years; AND
   - Is judged to be incapable of taking part in the assessment.

4. A student identified on the Administration Schedule as having and IEP or equivalent classification may be excluded from the assessment if:

   - The student is mainstreamed less than 50 percent of the time in academic subjects and is judged incapable of participating meaningfully in the assessment, OR
   - The IEP team or equivalent group has determined that the student is incapable of participating meaningfully in the assessment.

5. IEP/LEP students meeting the above criteria should be assessed if, in the judgment of school staff, they are capable of taking the assessment. **WHEN THERE IS DOUBT, INCLUDE THE STUDENT.**

If the staff decides to exclude any IEP or LEP students, please draw a line through the student's name making sure the name is still legible.

Next, prepare an Excluded Student Questionnaire for each of these students. Write the student's name on a Post-it note and affix it to the area labeled "Teacher of -" and give the questionnaire to the staff member most knowledgeable about the student. The Excluded Student Questionnaire should be returned by the day of the assessment and given to the NAEP Supervisor or Exercise Administrator who will complete the coding on the cover of the questionnaire.

5.5.3 Integrated Reading Performance Record (IRPR)

A subset of fourth-grade students who were assessed in reading were selected to participate in the Integrated Reading Performance Record, an individually administered audiotaped interview, developed to provide a richer context in which to assess a student's ability to read. During the interview, students were asked to read and answer questions about a story that was part of the NAEP assessment; students were also asked questions about their class work related to reading and materials they read independently.

After preassigning booklets, the supervisor identified those students who were assigned one of four booklets containing the preselected story. These students' names and assigned booklet ID numbers were recorded on a roster and those students not grade-eligible were excluded from selection for the study. Reading teachers of selected students were identified and received a notification letter as well a folder for each student sampled for the study. Students were asked to bring to the interview examples of their written classroom work as well as a book they recently read or were currently reading. Student names were removed from all interview materials and the materials were sent with the completed booklets in the school's shipment to NCS. A total of 1,421 IRPR interviews were conducted.

5.5.4 The Writing Portfolio Study

A subset of fourth- and eighth-grade students who were assessed in writing were selected to participate in a special project designed to provide information about the type of writing students do in school.

Portfolio folders were prepared and distributed to the primary writing teacher of selected students. Teachers were asked to assist students in selecting three pieces of writing that represented their best writing efforts, a range of writing tasks, and the use of writing process strategies. In addition, students were asked to write a letter to NAEP explaining why they chose their three pieces and what they like about each piece. Teachers were asked to fill out a brief questionnaire. The Writing Portfolio Roster was the central document used to keep track of every student sampled for the writing portfolio. Students' names were removed before their writing samples left the schools. A total of 1,828 portfolios were shipped to NCS.

5.5.5 Results of the Assessment

Of the 2,250 schools sampled for main NAEP assessments, 1,570 were assessed. In addition, 12 substitute school were assessed, for a total of 1,582 assessed schools. The assessed schools included 527 at age 9/grade 4, 587 at age 13/grade 8, and 468 at age 17/grade 12. The overall student participation rate was 87.5 percent. The 1992 participation rate of 93.4 percent among age 9/grade 4 students is slightly higher than the 1990 rate of 92.9 percent. At age 13/grade 8, the 1992 student participation rate of 88.8 percent remains comparable to past years, while the 80.8 percent response rate for age 17/grade 12 matches the rate obtained in 1990.

Of the 134,548 students sampled for the main assessments, 8.0 percent (10,825) were excluded by schools. Overall, 116,819 students were assessed across all three age/grade groups: 31,672 at age 9/grade 4; 42,591 at age 13/grade 8; and 42,556 at age 17/grade 12.

5.5.6 Assessment Questionnaires and Reports

Westat sent each school in the main NAEP assessments a School Characteristics and Policies Questionnaire a few weeks before the assessment was scheduled to be conducted. Likewise, supervisors prepared an Excluded Student Questionnaire for each sampled student the school deemed incapable of being assessed, following the NAEP exclusion criteria.

Selected teachers of fourth- and eighth-grade mathematics, fourth-grade reading, and eighth-grade writing were asked to fill out Teacher Questionnaires. The teachers asked to participate were the mathematics, reading, or writing teachers of the students who were assessed in those subject areas. The Teacher Questionnaire for grade 4 was combined into one form, since it is recognized that at grade 4 the mathematics and reading teacher would usually be the same person. For grade 8 there were two distinct questionnaires.

The supervisor requested that the Teacher Questionnaires distributed on the day of sampling be returned by the day of assessment, if possible. For those not returned on assessment day and those distributed after the assessment was completed, a postage-paid envelope was left with the school coordinator.

Table 5-4 shows the number of questionnaires that were distributed to main NAEP schools and the number shipped to NCS by the supervisors. These numbers do not include questionnaires mailed to NCS directly by the school coordinators.

Table 5-4
Completed Questionnaires Returned by NAEP Supervisors to NCS

Cohort	School Characteristics and Policies Questionnaires		Excluded Student Questionnaires		Teacher Questionnaires	
	Expected	Returned (%)	Expected	Returned (%)	Expected	Returned (%)
Age 9/Grade 4	527	517 (98%)	3412	3186 (93%)	1947	1756 (90%)
Age 13/Grade 8	587	553 (94%)	3963	3742 (94%)	2702	2325 (86%)
Age 17/Grade 12	468	439 (94%)	3450	3203 (93%)	—	—

5.6 FIELD MANAGEMENT

Three field managers assisted Westat home office project staff in monitoring the field work on the National Assessment. Assessment supervisors responsible for long-term trend assessments and a subset of assessment supervisors responsible for main assessments reported to these field managers, who in turn reported to Westat home office staff. The remaining main assessment supervisors reported directly to Westat's home office. Supervisors and field managers were required to report to their contacts at least once a week.

An automated management system was developed and maintained in Westat's home office. This system contained a record for each sampled school. A disposition code structure was developed to indicate the status of each school's participation (e.g., decision pending, school cooperating, school refusal, district refusal, school closed, school dropped - no age-eligibles, etc.). As the assessment supervisors reported the results of their contacts with district superintendents and individual schools, the disposition code was entered for each school. Disposition reports were then generated from the receipt system once a week so that home office staff could review the progress of securing cooperation from the sampled schools.

These reports were an invaluable tool for the sampling statisticians as well as for the field director and field management staff. They provided the statisticians with the information needed to determine whether the sample of schools was adequate to produce representative results. Based on the information contained in these reports, the sampling statisticians selected substitute schools to replace some of the noncooperating schools.

During the assessment activities, data from the School Worksheets on the number of students to be assessed, the number actually assessed, and the number absent were entered separately for long-term trend print- and tape-administered sessions. NCS provided Westat with similar data for main NAEP schools. In addition, data from questionnaire rosters on the number of Excluded Student Questionnaires, the number of Teacher Questionnaires expected and shipped to NCS, and whether the School Characteristics and Policies Questionnaire was shipped to NCS, were also entered into the receipt system. Data entered for the Integrated Reading Performance Record and the writing portfolio study included numbers of students sampled, interviews conducted, portfolios collected, and totals of shipments to NCS.

Weekly reports allowed the project staff to monitor the progress of the assessments both in terms of checking that the schools were assessed on schedule as well as assuring that a high response rate was achieved. The sampling statisticians used these reports to monitor the sample yield by school, PSU, and age/grade level.

Progress of the assessments was constantly monitored through telephone reports held between NAEP supervisors, field managers, and home office staff. During these phone conversations, the supervisors' schedules were reviewed, as well as any problems that the supervisors were experiencing.

The supervisors filled out a Work Schedule for a one- to two-week period, showing their whereabouts, so that they could be contacted if necessary. It also allowed field managers and project staff to review the supervisors' schedules and the distribution of work.

Progress of the field work was also monitored during quality control visits made to the field by Westat and ETS home office staff. The results of the quality control visits are discussed in the report on sampling and weighting procedures (Wallace & Rust, 1994).

Chapter 6

PROCESSING ASSESSMENT MATERIALS

Dianne Smrdel, Linda Reynolds, and Brad Thayer

National Computer Systems

6.1 OVERVIEW

This chapter describes the printing, distribution, receipt, processing and final disposition of materials for the National Assessment. The scope of the effort required by National Computer Systems (NCS) to process the materials is evidenced by the following:

- Prior to the assessment, approximately 24,800 bundles of assessment booklets were created and distributed to approximately 2,300 schools.
- For the approximately 117,000 students assessed at grades 4, 8, and 12 for the main assessments, about 135,000 assessment booklets and absent student forms and 17,000 questionnaires were received and processed, and about 2,053,000 constructed responses were professionally scored.
- For the approximately 36,000 students assessed for the fall, winter, and spring long-term trend assessments, about 40,000 assessment booklets and absent student forms and 3,300 questionnaires were received and processed and about 330,000 constructed responses were professionally scored.
- In all, approximately 3.3 million double-sided pages from test booklets and questionnaires were optically scanned.

Throughout the processing, the NCS Process Control System and Workflow Management System were used to track, audit, edit, and resolve characters of information. A quality control sample of characters of transcribed data was selected and compared to the actual responses in the assessment booklets.

The volume of collected data and the complexity of the National Assessment processing design, with its spiraled distribution of booklets, required the enhancement and implementation of flexible, innovatively designed processing programs and a sophisticated Process Control System. This system, developed for the 1990 assessments, allowed an integration of data entry and workflow management systems, including carefully planned and delineated editing, quality control, and auditing procedures.

The magnitude of the processing effort is apparent when considering that the activities described in this chapter were completed within a short time frame, and that an estimated accuracy rate of fewer than five errors for every 10,000 characters of information was achieved.

Several major changes in materials processing were made from 1990, including the conversion of all documents to scannable form, the tailoring of shipments to the individual size and requirements of schools, and the reorganization of the process flow to conduct constructed-response scoring after all machine scoring and data verification processes were complete, allowing NCS to provide Westat and ETS with demographic and cognitive data at an earlier date.

6.2 PROCESS CONTROL SYSTEM

NCS maintains a Process Control System consisting of numerous specialized programs and processes to accommodate the unique demands of concurrent assessment processing and a unified ETS/NCS system integration. The Process Control System, which was developed for the 1990 assessment, was necessary to maintaining control of all shipments of materials to the field and from the field, and of any work in progress. The system is a unique combination of several reporting systems currently in use at NCS, along with some application-specific processes. These systems are the Workflow Management System, the Bundle Assembly Quality Control System, the Outbound Mail Management System, and the On-line Inventory Control system. Data were collected from these systems and recorded in the file called the "NAEP Process Control System." Additional information was entered directly into the Process Control System.

6.3 WORKFLOW MANAGEMENT SYSTEM

The functions of the Workflow Management System are to keep track of where the production work is and where it should be and to collect data for status reporting, forecasting, and other ancillary subsystems. The primary purpose of the Workflow Management System is to analyze the current workload by project across all workstations.

The data processing and control systems are determined to a large extent by the type of documents processed. For the National Assessment, only machine-scannable assessment booklets and answer documents were used to collect student responses. The five questionnaires that were used to collect data about school characteristics, teachers associated with sampled students, and students excluded from the assessment were also scannable documents.

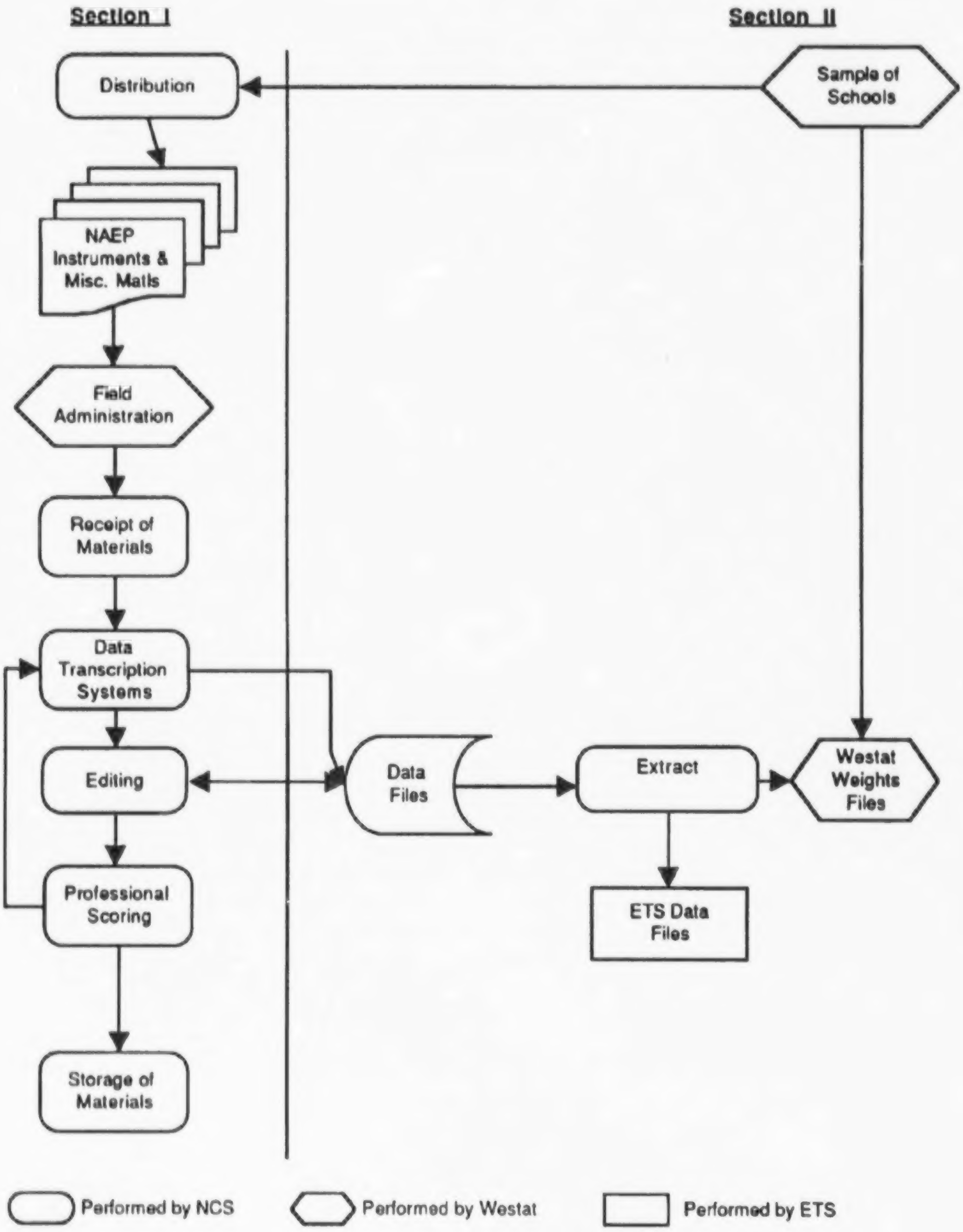
6.4 PROCESS FLOW OF NAEP MATERIALS AND DATABASE CREATION

Figure 6-1 shows the conceptual framework of processes that were used for the National Assessment materials.

Section I of Figure 6-1 depicts the flow of NAEP's printed materials. Information from the Administration Schedule and Packing List was used to control the processing of materials. The figure follows the path of each assessment instrument—Student Test Booklets, School

Figure 6-1

Data Flow Overview, 1992 National Assessment of Educational Progress


Section I
Section II
Distribution
Sample of Schools
NAEP Instruments & Misc. Matls
Field Administration
Receipt of Materials
Data Transcription Systems
Editing
Data Files
Extract
Westat Weights Files
Professional Scoring
ETS Data Files
Storage of Materials
Performed by NCS
Performed by Westat
Performed by ETS

Characteristics and Policies Questionnaires, Teacher Questionnaires, Excluded Student Questionnaires, Packing List, and Administration Schedules—as they were tracked through the appropriate processes that resulted in the final integrated NAEP database.

The remainder of this chapter provides an overview of the materials processing activities as shown in Section I of Figure 6-1. Section II of Figure 6-1 depicts the evolution of the NAEP/NCS database from the transcribed data to the final files, provided to Westat for creation of weights and to ETS for analysis and reporting.

The 1992 NAEP data collection resulted in six classes of data files (student, school, teacher, excluded student, sampling weight, and item information files). The structure and internal data format of the 1992 NAEP database was a continuation of the integrated design originally developed by ETS in 1983.

6.5 MATERIALS DISTRIBUTION

The use of bar code technology in document control was introduced to NAEP by NCS in the 1990 assessment; its use continued in 1992. A bar code was applied to the front cover of each document. The bar code consisted of a two-digit booklet number, a five-digit sequential number, and a check digit.

The booklets were spiraled into unique bundles consisting of 11 booklets in a set pattern. A header sheet was attached to each bundle that indicated the assessment type, bundle type, bundle number, and a list of the booklet types to be included in the bundle.

The bundle numbers on the header sheet were created to identify the type of bundle. Before they were distributed, all bundles were passed under a scanner programmed to interpret the bar code and the file of scanned barcodes was transferred to the mainframe. A computer program compared the bundle type expected to the one actually scanned after the header and verified that there were 11 booklets in each bundle. Any discrepancies were printed on an error listing forwarded to the Packaging Department, where the error was corrected and the bundle was again read into the system for another quality control check. This process was repeated until all bundles were correct. The bundles were shrink-wrapped in clear plastic. The bundles were then ready for assignment and distribution.

When packing lists for distribution of materials were created from the Materials Distribution System, a second and more detailed bundle slip was produced. This bundle slip indicated the same information as the slip wrapped with the bundle, plus the school number and the ID numbers of the booklets within that bundle. This allowed the assessment administrators to pre-assign booklets for their sessions.

Each school conducted at least one session; some conducted more than one. The materials needed for a school to conduct all of its sessions were sent to the supervisors in one shipment. In 1990, each session's materials had been shipped independently. Although this change in shipment practice eliminated the option to pre-assemble many materials, it did cause less confusion in the field. The following materials were shipped:

Bundle(s) of 11 assessment booklets (based on sample count)
Scientific calculators (grade 8 mathematics) per bundle of booklets
Four-function calculators (grade 4 mathematics) per bundle of booklets
Protractors
Cassette tape for estimation booklets
Digital timer
Calculator posters
Mathematics posters
Tape recorder with batteries
Rulers
Sets of geometric shapes
Pad of appointment cards
Return postage-paid labels
Post-it note pads
Shipping tape
Excluded Student Questionnaires
Teacher Questionnaires
Roster of Questionnaires
Assessment Notifications
Pre-addressed envelopes
School Characteristics and Policies Questionnaires
Pre-addressed box

Shipments were sent according to the week of assessment. Some supervisors found they needed extra quantities of materials (e.g., more Excluded Student Questionnaires or Teacher Questionnaires) and calls were received requesting these additional materials.

Aiding in the security of the shipments was the decision to send all shipments, whenever possible, through the Airborne delivery system. NCS is connected to the Airborne system through computer link, thus expediting tracing of any misdirected shipments. This system provides the date and time of delivery as well as the name of the person who signed for the shipment. All shipments were recorded in the Airborne Libra system. If a shipment had to be sent by United Parcel Service or the U.S. Postal Service, this information was also recorded and transferred to the mainframe.

6.6 PROCESSING ASSESSMENT MATERIAL

The materials from each session were to be returned to NCS in the same box in which they were originally mailed. It was the responsibility of the exercise administrator to repackage the items in the proper order, complete all paperwork and return the shipment through the U.S. Postal Service, using the postage-paid label provided.

Given the quantity of materials arriving at NCS in a short period of time, it was necessary to devise a system that would quickly acknowledge receipt of a school's material. A label applied to the outside of the box by the NCS packaging department contained a bar code that indicated the school number and the project number. When the shipment arrived at NCS, the bar code was read and the shipment forwarded to the receiving area. The file was then transferred to the mainframe through a PC link and a computer program was used to apply the shipment receipt date to the appropriate school within the Process Control System. This provided the current status of shipments received regardless of any processing backlog. This information was then transferred electronically to Westat. The status of the administration was checked and in cases of a missing shipment, a trace was initiated.

Receiving personnel checked the shipment to verify that the contents of the box were complete and accurate, matching the school and session indicated on the label.

The materials were checked against the Packing List to verify that all materials were returned. If any discrepancies were found, an alert was issued. If all assessment instruments were returned, processing continued.

Each booklet and Excluded Student Questionnaire was verified against the Administration Schedule. All counts of booklets returned were verified and the information on the front cover of the booklets was matched to that on the Administration Schedule. If any discrepancy was discovered, an alert was issued.

After the contents of the shipment had been identified and verified, the information from the Administration Schedule was entered into the Process Control System. That information included school number, session code, counts of the number of students in the original sample, students in the supplemental sample, students in the total sample, students withdrawn, students excluded, students to be assessed, students absent, students originally assessed, students assessed in makeup sessions, and the total number of students assessed. If a makeup session was expected, an information alert was issued to facilitate tracking. The control counts were used by NCS for verification of processing counts. This information was also transferred electronically to Westat weekly to be used to produce participation statistics for the states.

If quantities and individual information matched, the booklets were organized into work units and batched for processing. The processing flow was changed in 1992, resulting in the completion of the machine scoring prior to the constructed-response scoring. Each batch, consisting of multiple sessions, was assigned a unique batch number. The batch number was entered on the Workflow Management System, facilitating the internal tracking of the session and allowing departmental resource planning. A scannable session header, included in the shipment from the school, was coded with the session code and placed on top of the stack of documents. All student documents were forwarded to machine scanning functions. Control documents were forwarded to appropriate record filing systems.

The Excluded Student Questionnaires and Teacher Questionnaires were compared to the Roster of Questionnaires and the Administration Schedule to verify demographic information. Some questionnaires may not have been available for return with the shipment. These were returned to NCS at a later date in an envelope provided for that purpose. If the Excluded Student Questionnaire was not returned with the shipment of booklets, a record containing all demographic information on that student from the Administration Schedule was entered into the Process Control System. If the questionnaire were subsequently returned, this record was deleted. Otherwise, the record was provided to Westat for use in the weighting process.

Each School Characteristics and Policies Questionnaire was compared with the Roster of Questionnaires and the school number was verified to match all other materials in the shipment. As with the other questionnaires, this document may not have been returned with the shipment and could also be returned in the supplemental envelope. There was no additional effort made to collect or report information on unreturned school questionnaires.

All assessed and absent students were assigned a test booklet. To indicate an absence, the "A" bubble in the Administration Code column on the front cover of the booklet was gridded. The booklet was then processed with assessed student booklets to maintain session integrity.

The Packing List was used by the supervisors to account for all materials received from and returned to NCS. Any discrepancies in quantities received or returned to NCS were indicated. Also indicated was whether a makeup session was to be held, the date of scheduled makeup, the number of students involved, and the quantities of materials being held for later return.

The Administration Schedule contained the demographic characteristics of the students selected for the assessment. This information included gender, race/ethnicity, birth date, and IEP/LEP indicators. The booklet number of the student selected was recorded on the Administration Schedule during the assessment process, and the demographic information was transferred to the booklet covers by either the student or the assessment administrator.

The demographics of the sampled students who did not participate in the assessment (exclusions and absentees) were provided to Westat to be used to adjust the sampling weights of the students who did participate. The excluded student information was obtained from the Excluded Student Questionnaire or provided on a file for those questionnaires not returned to NCS. The absent student information was taken from the front cover of the booklet that was assigned prior to the start of the assessment. This procedure eliminated the need for an additional form for absent students.

For each type of questionnaire, two numbers were entered on the Rosters of Questionnaires: number of questionnaires expected and number actually received. The Packing List, Administration Schedule, and Roster of Questionnaires were forwarded to the operations coordinator and filed by school within state for future reference.

6.7 DATA TRANSCRIPTION SYSTEMS

The transcription of the student response data into machine-readable form was achieved through the use of three separate systems: data entry (scanning), validation (pre-edit), and resolution.

6.7.1 Data Entry

The data entry process was the first point at which booklet-level data were input to the computer system. As all documents used in the 1992 assessment were scannable documents, the data were collected using NCS optical scanning equipment. The data were then edited and questionable data were resolved before further processing.

To ensure data integrity, edit rules were applied to each scanned data field. This procedure validated each field and reported all problems for subsequent resolution. After each field was examined and corrected, the edit rules were re-applied for final verification.

6.7.2 Scanning

After the initial manual verification, the scannable documents were transported to a slitting area where the folded and stapled spine was removed from each document. Scanning operations were performed by NCS's HPS Optical Scanning equipment. The optical scanning devices and software used at NCS permit a complete mix of NAEP scannable materials to be scanned with no special grouping requirements. However, for manageability and tracking purposes, student documents, Excluded Student Questionnaires and Teacher Questionnaires were batched separately. The bar code identification numbers used to maintain process control were decoded and transcribed to the NAEP computerized data file.

The scanning program is a table-driven software process that uses standard routines and application-specific tables to identify and define the documents and formats to be processed. When a booklet cover is scanned, the program uses the booklet number to determine the sequence of pages and the formats to be processed. By reading the booklet cover, the program recognizes which pages should follow and in what order.

The scanning program wrote four types of data records into the data set: a batch header record containing information coded onto the batch header sheet by receipt processing staff; a session header record containing information coded onto the session batch header sheet by receipt processing staff; a data record containing all of the translated marked ovals from all pages in a booklet; and a dummy data record, serving as a place holder in the file for a booklet with an unreadable cover sheet. The document code was written in the same location on all records to distinguish them by type.

The following coding rules were used:

- The data values from the booklet covers and scorer identification fields were coded as numeric data.
- Unmarked fields were coded as blanks and processing staff were alerted to missing or uncoded critical data.
- Fields that had multiple marks were coded as asterisks (*).

- The data values for the item responses and scores were returned as numeric codes.

- The multiple-choice, single-response format items were assigned codes depending on the position of the response alternative; that is, the first choice was assigned a 1, the second a 2, and so forth.

- The circle-all-that-apply items were given as many data fields as response alternatives; the marked choices are coded as 1 and the unmarked choices as blanks.

- The fields from unreadable pages were coded with an X as a flag for resolution staff to correct.

6.8 DATA VALIDATION

The data entry and resolution system used for the National Assessment is able to process materials submitted from both scannable and nonscannable media simultaneously for three age groups, three assessment types, and five questionnaires. The use of batch identification codes—comprising the school and session codes as well as the batch sequence numbers for suspect record identification—facilitated the management of the system and correction of incorrectly gridded or keyed information.

As the program processed each data record, it first read the booklet number and checked it against the batch session code for appropriate session type. Any mismatch was recorded on the error log and processing continued. The booklet number was compared against the first two digits of the student identification number. If they disagreed because of improper bar coding, a message was written to the error log. The remaining booklet cover fields were then read and validated for the correct range of values. The school codes had to be identical to those on the Process Control System record and the grade code had to be either 4 or 8. All data values that were out of range were read as is, but flagged as suspect. All data fields that were read as asterisks were recorded on the edit log.

Document definition files describe each document as a series of blocks that are described as a series of items. The blocks in a document were traversed in the order that they appear on the document. Each block's fields were validated during this process. If a document contained suspect fields, the cover information was recorded on the edit log with a description of the suspect data. Some fields (e.g., AGE or DOB) required special types of edits. These fields were identified in the document definition fields, and a subroutine was invoked to handle these cases.

The program next cycled through the data area corresponding to the item blocks. The task of translating, validating, and reporting errors for each data field in each block was performed by a routine that required only the block identification code and the string of input data. This routine had access to a block definition file that had the number of fields to be processed for each block and the field type (alphabetic or numeric), the field width in the data record, and the valid range of values for each field. The routine processed each field in sequential order, performing the necessary translation, validation, and reporting tasks.

The first of these tasks checked for the presence of blanks or asterisks in a critical field. These were recorded on the edit log and processing continued with the next field. No action was taken on blank-filled fields for multiple-choice items since that code indicated a nonresponse. The field was validated for range of response, recording anything outside of that range to the edit log. The item type code was used by the program to make a further distinction among constructed-response item scores and other numeric data fields. Moving the translated and edited data field into the output buffer was the last task performed in this phase of processing.

The completed string of data was written to the data file when the entire document had been processed. Then, when the next session header record was encountered, the program repeated the same set of processes for that session. The program closed the data set and generated an edit listing when it encountered the end of a file.

Accuracy checks were performed on each batch processed. Every 500th document of each booklet form was printed in its entirety, with a minimum of one document type per batch. This record was checked, item by item, with the source document for errors.

6.9 EDITING

Quality procedures and software throughout the system ensure that the NAEP data are correct. The initial editing that took place during the receipt control process included verification of the schools and sessions. Receipt control personnel checked that all student documents on the Administration Schedule were undamaged and assembled correctly. The machine edits performed during data capture verified that each sheet of each document was present and that each field had an appropriate value. All batches entered into the system were edited for errors.

Data editing occurred after these checks and consisted of a computerized edit review of each respondent's document and the clerical edits necessary to make corrections based upon the computer edit. This data editing step was repeated until all data were correct.

The first phase of data editing was designed to ensure that all documents were present. A computerized edit list was produced after NAEP documents were scanned and with the supporting documentation sent from the field the edit function was performed. The hard copy edit list contained all the vital statistics about the batch and each school and session within the batch, such as the number of students, school code, type of document, assessment code, error rates, suspect cases, and record serial numbers. Using these inputs, the data editor verified that the batch had been assembled correctly, each school number was correct, and all student documents within each session were present.

During data entry, counts of documents processed by type were generated. These counts were checked against the Administration Schedule counts entered into the Process Control

System during the receiving process. The number of assessed and absent students processed had to match the number of used booklets indicated on the Process Control System.

The second phase of data editing was carried out by an experienced editing staff using a predetermined set of rules to review the field errors and record corrections to be made to the student data file. The same computerized edit list used in the first phase was also used to perform this function.

The editing staff made corrections using the edit log prepared by the computer and the actual source document listed on the edit log. The corrections were identified by batch sequence numbers and field name for suspect record and field identification. The edit log indicated the current composition of the field. This particular piece of information was then visually checked against the NAEP source document by the editing staff for double grids, erasures, smudge marks or omitted items that were flagged. Each flagged item was handled in one of the following ways:

- *Correctable Error:* If the error could be corrected by the editing staff, according to the editing specifications, the corrections were indicated on the edit listing.
- *Field Correctable:* If an error was not correctable according to the specifications, an alert was issued to the operations coordinator for resolution. Once the correct information was obtained, the correction was indicated on the edit listing.
- *Noncorrectable Error:* If a suspected error was found to be correct as stated, and no alteration was possible according to source documents and specifications, the programs were tailored to allow this information to be accepted into the data record and no corrective action was taken.

These corrections were noted on the edit list. When the entire batch of sessions was resolved, the list was forwarded to the key entry staff. The corrections were entered and verified through the Falcon system. When all corrections were entered and verified for a batch, an extract program was run to pull the correction records to a mainframe data set.

The post-edit program was initiated next. This program applied the corrections to the specified records and once again applied the error criteria to all records. If there were further errors, another edit list was printed and the cycle began again.

When the edit process had produced an error-free file, the booklet ID number was posted to the NAEP tracking file by school and sessions. This allowed for an accumulation process to accurately measure the number of documents processed for a session within a school and the number of documents processed by form. The posting of booklet IDs also ensured that a booklet ID was not processed more than once. These data allowed the progress of the assessment to be monitored and reported on the status report.

At this point, a job was automatically submitted to produce the NAEP scoring sheets for the batch. The program also selected the records to be scored by a second reader for reliability. These sheets were printed, matched with the original documents, and forwarded to the NAEP scoring area. (The professional scoring procedures are described in Chapter 7.)

Once all documents for a batch had been scored, the sheets were batched and submitted to scanning. A series of edits were run to verify the information on these sheets. The scorer identification fields were processed at this point and certain checks were made. The routine validated the score range and did not permit a blank field. If no score was indicated or the score was out of range, the disparity was noted on the edit log.

These error logs were returned to the scoring groups for resolution and the corrections were entered directly to the files. The edit process was repeated until the file was error free.

As a final quality control check, ETS identified a random sample of each booklet type from the master student file (see Chapter 8). The designated documents and scoring sheets were located, removed from storage and forwarded to ETS for quality control. On completion of quality control processing, the booklets were returned to NCS for return to storage.

6.10 QUESTIONNAIRES

The questionnaires were received either with the session shipment or in a later shipment. Once the questionnaires were verified with the roster, they were accumulated by the receiving clerks. The School Characteristics and Policies Questionnaires, Teacher Questionnaires, and Excluded Student Questionnaires were batched and sent to scanning at regular intervals. Every effort was made to keep current on all forms, both to ensure the processing of all documents for a session and to deliver all data at the same time.

All documents, regardless of method of entry, were run through the process of error identification and resolution.

6.11 MERGING OF STUDENT DATA

At the completion of the scoring and verification of the constructed responses, the complete records for students were merged. This merge included the machine-scanned data, the scores to the constructed responses, and the responses from the estimation booklets. Verification of complete student records was conducted prior to the delivery of the data files.

6.12 STORAGE OF DOCUMENTS

Once the editing process had been successfully completed on the batches, they were sent to the NCS warehouse for storage. The storage location of all documents was recorded on the inventory control system and stored for later retrieval. Unused materials were sent to temporary storage until the completion of the assessment and acceptance of the data files, at which time they were destroyed.

Chapter 7

PROFESSIONAL SCORING

Andrew S. Latham

Educational Testing Service

Brad Thayer, Lavonne Mohn, Brent Studer, Connie Smith-Hora, and Robert Neiers

National Computer Systems

The 1992 assessment included a variety of constructed-response tasks—that is, questions that required students to produce a response rather than to select the correct answer from a series of options. Constructed-response questions were included in the main and long-term trend assessments in mathematics, reading, and writing, as well as in the Trial State assessments in fourth-grade reading and in fourth- and eighth-grade mathematics. The responses to these questions were professionally scored by teams of trained readers.

The 1992 main, long-term trend, and Trial State assessments included the following numbers of constructed-response questions:

Table 7-1
Numbers of Constructed-Response Questions

Subject Area	1992 Main and Trial State Assessments				
	Age 9/ Grade 4	Age 13/ Grade 8	Age 17/ Grade 12	Trial State Grade 4	Trial State Grade 8
Mathematics	71	75	77	59	66
Reading	43	79	85	43	N/A
Writing	9	11	12	N/A	N/A

Subject Area	1992 Long-term Trend Assessments		
	Age 9	Age 13	Age 17
Mathematics	44	27	57
Reading	5	8	9
Writing	6	6	6

The following sections summarize the scoring of constructed-response questions in the main, long-term trend, and Trial State assessments. The last section presents scoring information about two special studies: the Integrated Reading Performance Record and the Writing Portfolio. A more detailed discussion can be found in the *Report of Processing and Professional Scoring Activities, NAEP - 1991-92* (Kennel, Mohn, Reynolds, Smrdel, Thayer, & Moyer, 1992).

7.1 MAIN AND TRIAL STATE MATHEMATICS ASSESSMENTS

7.1.1 Characteristics of the Scoring Guides

The mathematics portion of the 1992 National and Trial State assessments included 161 discrete constructed-response questions designed to measure different aspects of students' understanding of mathematics. Some of these questions were administered at more than one grade level. In addition, a special study booklet at each grade contained estimation and innovative problem-solving questions which were paced by an audiotape. Scoring guides for the 161 constructed-response questions provided correct responses and solutions for each problem, alternative methods of solutions that could be used by students to solve the problems correctly, and certain incorrect student responses that indicated particular errors or misunderstandings.

The constructed-response mathematics questions in the 1992 NAEP assessment measured five major content areas and three levels of mathematical ability that required students to demonstrate mathematical understanding and problem-solving techniques at various levels of sophistication. The items required the scoring of computational answers, patterns, tables and charts, geometric figures, graphs, and brief statements of explanation or justification. To this end, each scoring guide was prepared and refined as an integral part of the item development process. At each step of the two-year NAEP item development and review process, the scoring guide for a given constructed-response question underwent scrutiny and discussion and was revised as necessary. In addition, each scoring guide was subject to a final review and revision, if necessary, as part of the training process. While selecting sample responses for use in training readers, NAEP staff examined a variety of student responses to each question to determine the appropriateness of the scoring guide and to select illustrative student responses that were to be included in the packet of materials that was used to train readers.

The constructed-response questions were divided into two categories of "short" and "extended" responses. The short constructed-response questions required short answers that were dichotomously scored as either right or wrong, with several types of wrong answers classified to provide additional data about students' mathematics understanding. The extended constructed-response questions required more complex answers, which were scored on a scale of 1 to 5 to allow for partial credit. The score-point outlines for short and extended constructed responses are given in Figure 7-1.

Figure 7-1

Outline for Scoring of Mathematics Constructed Responses

Short Constructed Responses

Score	Definition
	Scores Awarded for Full Credit and Counted as Correct Responses:
8	This score indicated a correct answer in the detail required for commonly used methods of solution.
7	This score indicated a correct answer in the detail required for alternate solutions.
	Scores Awarded for Partial Solutions but Counted as Incorrect Responses:
5,6	These scores were given to responses that were correct to a point but were either incomplete or contained some error or irrelevant information.
	Scores Awarded No Credit and Counted as Incorrect Responses:
2,3,4	These scores indicated an incorrect response to the questions that clearly reflected a student's misunderstanding of the concept being measured or a commonly given incorrect response.
1	This score was given to responses that were incorrect but could not be categorized with typical incorrect responses.

Extended Constructed Responses

4	This score was given to responses which demonstrated EXTENDED mathematical reasoning.
3	This score was given to responses which demonstrated SATISFACTORY mathematical reasoning.
2	This score was given to responses which demonstrated PARTIAL mathematical reasoning.
1	This score was given to responses which demonstrated MINIMAL mathematical reasoning.
9	This score was given to responses which demonstrated INCORRECT mathematical reasoning.

A score of 0 was given to questions for which there was no response.

A score of 9 was given to responses that were indecipherable, irrelevant, or contained a statement to the effect that the student did not know how to do the task.

7.1.2 Training

The readers were organized into 11 teams, each containing approximately 9 readers and a team leader. In late January, 1992, before the training process began, ETS test development staff and the team leaders prepared training sets (sets of sample responses to accompany the scoring guides) and refined the scoring guides. ETS mathematics specialists conducted the training of the readers with assistance from the team leaders. Training involved explaining each item and its scoring guide and discussing responses that were representative of the various score points in the guide. When this was completed, the readers scored and discussed 5 to 20 randomly selected "practice papers" for each item, depending on the complexity of the item. The purpose of the training was to familiarize the readers with the scoring guides and teach them how to apply the scoring criteria to a wide range of student responses in a consistent and reliable manner. Following the group training, each reader on the teams scored all of the constructed responses in each of approximately 12 bundles of booklets, each of which contained an average of 54 booklets. During this practice, discussion sessions were held to review and clarify responses for which subjective judgments were required. In some cases the scoring guides were revised to track common incorrect answers that had not appeared in the original sample sets. The initial training was completed in one week.

7.1.3 Scoring

Materials from the 1992 main and trial state assessments were scored at the same time and by the same readers. To determine interreader reliability, 20 percent of the mathematics constructed responses were scored by a second reader. The team leaders monitored the scoring of all members of the team by reading individual responses and the corresponding scores given by the reader. Individual and/or group discussions were initiated when the team leader found evidence of a reader or readers having difficulty scoring certain items.

The percent of agreement between readers was high, with an average agreement of 94.1 percent across all constructed-response questions. The percent of exact agreement between readers for the 11 extended constructed-response tasks ranged from 69.3 percent to 90.5 percent, with an average of 81.1 percent. For the 82 short constructed-response questions, the percent of exact agreement between readers ranged from 81.9 to 99.2 percent, with an average of 96.2 percent. This reliability information was also used by the team leaders in monitoring the capabilities of all readers and the uniformity of scoring across readers.

Because the reliability scoring was done on separate scoring sheets, all reliability scoring was "blind," or uninfluenced by any score already given. The reliability scoring for each batch of booklets was completed by a team other than the one that did the full scoring. In this way, the full scoring and the reliability scoring was spread across all teams so that readers were compared against all other readers.

7.2 MAIN AND TRIAL STATE READING ASSESSMENTS

7.2.1 Characteristics of the Scoring Guides

The 1992 main NAEP reading assessment for grades 4, 8, and 12 and Trial State reading assessment for grade 4 contained a total of 170 discrete constructed-response questions, 135 of which were short constructed-response questions and 35 of which were extended-response questions. Included in the distribution of items for grades 8 and 12 were the NAEP Reader blocks. The NAEP Reader gave students the opportunity to choose one story from a set of seven and respond to 12 questions about that story.

The reading scoring guides were developed using the same rigorous standards used for the mathematics guides. The reading assessment framework encompassed two reading purposes at grade 4: reading for literary experience and reading for information. A third purpose, reading to perform a task, was added at grades 8 and 12. To best account for these varied purposes, NAEP rated the short constructed responses on a right/wrong basis, and the extended constructed responses on a scale of 1 to 4, which allowed for partial credit.

Figure 7-2
Outline for Scoring Constructed Responses for Reading

Short Constructed Responses

Score	Definition
4	This score was given to responses that demonstrated ACCEPTABLE comprehension.
1	This score was given to responses that demonstrated UNACCEPTABLE comprehension.

Extended Constructed Responses

4	This score was given to responses that demonstrated EXTENSIVE comprehension.
3	This score was given to responses that demonstrated ESSENTIAL comprehension.
2	This score was given to responses that demonstrated PARTIAL comprehension.
1	This score was given to responses that demonstrated UNSATISFACTORY comprehension.

A score of 0 was given to questions for which there was no response.

A score of 9 was given to responses that were indecipherable, irrelevant, or contained a statement to the effect that the student did not know how to do the task.

7.2.2 Training

The training procedures used for the reading assessments closely resembled those described for the mathematics assessment. Working with NCS staff, ETS content area specialists refined the scoring guides and selected illustrative sample papers.

Sixty readers organized into four teams were trained to score the grade 4 main and Trial State short constructed responses for reading; 11 readers were trained to score the grade 8 and grade 8/12 main assessment overlap reading responses; and 10 readers were trained to score the grade 12 main reading responses.

Because of the large number of constructed responses that had to be scored, ETS conducted the training for the four teams scoring the grade 4 items differently from the training for the teams scoring the grades 8 and 12 items. The training for the grades 8 and 12 teams was conducted over the course of one week by the same trainer on all applicable items. The training for the grade 4 teams was conducted over the course of one week by rotating trainers among the four teams of readers. Each grade 4 trainer specialized in two blocks of questions and trained each of the four teams on these blocks.

The overall percentage of agreement between readers for the national reliability samples at each of the three grades assessed was 89 percent at grade 4, 86 percent at grade 8, and 88 percent at grade 12.

7.2.3 NAEP Reader

The NAEP Reader portion of the assessment consisted of two NAEP Reader books, one each for grades 8 and 12, each of which contained seven different stories. Although students chose which story they wanted to read, all students at each grade were asked the identical set of nine short and three extended constructed-response questions. ETS subject area specialists created scoring templates for each item. For each of the possible 14 NAEP reader stories, unique examples of responses accompanied the criteria listed in the standards.

Training

Because of the frequency with which certain of the NAEP Reader stories were chosen and the need to expedite the training process on such a large number of unique scoring criteria, the team of nine readers was divided into two teams of four and five readers. At grade 8, five of the seven stories were chosen much more frequently than the other two stories. At grade 12, three of the seven stories were chosen more frequently. The training was divided between these two groups of readers to accommodate the eight stories selected most frequently by students.

Scoring

The constructed-response scoring standards for the NAEP Reader stories used the same approach to scoring as was used with all other reading items. The short constructed-response questions were scored using the dichotomous scale of acceptable versus unacceptable responses, while the extended-response questions were scored using the four-point scale.

7.2.4 Integrated Reading Performance Record

The 1992 NAEP assessment also included the Integrated Reading Performance Record, a special study of fourth-grade students' oral reading proficiency. Students participating in the study were audiotaped as they responded to an interview, read a passage from the written portion of the assessment aloud, and answered questions about the passage. The tape-recorded interview was transcribed by NCS communication center staff and the transcript of the student interview was included in each student's folder along with the tape itself. The scoring of this special study project was broken into three discrete scoring parts, discussed below:

- Interview Analysis;
- Miscue Analysis; and
- Reading Comprehension and Fluency Evaluation Scoring.

Interview Analysis

The Interview Analysis consisted of categorizing students' oral responses to questions regarding various personal and school-related reading activities. This categorization of the oral responses captured information about these reading activities in order to provide teachers, administrators, and legislators with data to aid educational decision-making.

Training. The 16 scorers were divided into two teams of 8 readers, with one team leader for each team. Each reader received a scoring template which outlined the response categories for each of the questions. The training coordinator reviewed the scoring template and discussed each of the response categories, describing the types of responses which might fall into these categories and elaborating on the rationale for the particular question. Prior to beginning scoring, the teams, as a group, practice-scored a number of student interviews, then moved into the scoring process, closely monitored by the team leaders.

Scoring. The scoring template for the Interview Analysis outlined the various interview prompts and the range of possible response categories. Using this template as their guide and the interview transcript as the written representation of the students' responses, the readers categorized all student responses on the Interview Analysis Coding sheet.

Miscue Analysis

The Miscue Analysis portion of the Integrated Reading Performance Record involved listening to students' oral readings of a passage and categorizing any problems associated with these readings. For example, if the student paused at inappropriate places in a sentence, this error was identified and coded by the scorers. Other typical reading errors included the repetition, omission, or mispronunciation of words. The categorization of oral reading miscues was intended to provide more detailed information about the types of problems students have in oral reading situations, particularly at the level of word recognition and pronunciation.

Training. The 24 readers were divided into three teams of 8 readers, with one team leader for each team. The training coordinator conducted this portion of the training process by having all readers listen to individual student readings and by describing the types of responses which might fall into the main categories of substitution, insertion, and omission. Prior to beginning scoring, the teams practice-scored a number of student tapes under the constant supervision and guidance of the team leaders.

Scoring. The Miscue Analysis portion of the study was coded on a typed script of the reading passage. Then this information was transferred to a key-entered Miscue Coding Sheet.

Reading Comprehension and Fluency Evaluation Scoring

The Reading Comprehension and Fluency Evaluation Scoring of the Integrated Reading Performance Record served two purposes. The first was to gain some sense of whether students' oral responses to three reading comprehension questions would differ markedly from their own written responses to the same questions. The second was to assign an overall reading fluency rating to their oral reading of the passage. These scorings were intended to provide descriptive information on reading skills as they pertain to the oral response versus the written response and reading skills as they pertain to appropriate or meaningful oral interpretations of the language of the text.

Training. The preparation for training for the Reading Comprehension portion of the study was conducted by the same team leader who had assisted in the training during the written reading comprehension assessment. This preparation for training consisted of organizing the same materials that had been used to train readers originally.

The preparation for training for the Fluency Evaluation Scoring portion of the study was conducted by the ETS subject area specialist. Scorers practiced a holistic scoring approach to reading fluency by simply assigning an overall fluency evaluation based on the scorer's trained impression of reading ability, cadence, etc. Prior to scoring, the team practice-scored a number of student tapes. Readers then proceeded with the scoring process, always closely monitored by the team leader.

Scoring. Both the Reading Comprehension and Fluency Evaluation portions of the study were scored on a key-entered form, with space for the three reading comprehension scores and the one holistic fluency score. A minimum of twenty percent of the folders were scored by a second reader to obtain interreader reliability statistics.

7.3 MAIN WRITING ASSESSMENT

7.3.1 Characteristics of the Scoring Guides

The 1992 NAEP writing assessment for grades 4, 8, and 12 contained a total of 22 discrete constructed-response tasks. Unlike the mathematics and reading assessments, the writing assessment had no trial state component, and no multiple-choice or short constructed-response questions. Students wrote either two 25-minute papers or one 50-minute paper, all of which were scored on a six-point primary trait scale (Figure 7-3).

Figure 7-3

Outline for Scoring Writing Responses

Primary Trait

Score	Definition
6	This score was given to EXTENSIVELY ELABORATED responses.
5	This score was given to ELABORATED responses.
4	This score was given to DEVELOPED responses.
3	This score was given to MINIMALLY DEVELOPED responses.
2	This score was given to UNDEVELOPED responses.
1	This score was given to RESPONSES TO THE TOPIC (but not the task).

A score of 0 was given to questions for which there was no response.

A score of 9 was given to responses that were indecipherable, irrelevant, or contained a statement to the effect that the student did not know how to do the task.

The 1992 writing framework focused on three purposes for writing: informative, persuasive, and narrative. Informative papers communicate detailed information, persuasive papers convince the reader of a point of view, and narrative papers create stories from actual or imagined events. In addition to the primary trait tasks, students were given a blank page which they were encouraged to use to help them plan their writing. Many students took advantage of this option. Their pre-writing was classified into seven categories which were not mutually exclusive:

Category	Definition
1	Copy of prompt
2	Unrelated notes, drawings, or comments
3	Lists, outlines, or notes
4	Diagram
5	False starts or recopying
6	Different version
7	First draft

7.3.2 Training

As with the mathematics and reading assessments, the formal training for the writing assessment was conducted by ETS subject area specialists. In addition, to help prepare for the special challenges posed by the six-point primary trait scale to assess student writing, team leaders compiled two sets of calibration packets of 50 to 60 papers to be used subsequent to the formal training process. These calibration packets refocused readers' attentions on specific criteria for each task and gauged the level of reader understanding of the various items prior to the actual scoring.

7.3.3 Scoring

A 25 percent second-scoring was conducted to determine interreader reliability. The broad spectrum of student responses, which could range from a few sentences to a half-dozen pages, coupled with the ambitious 6-point scoring guides, made accurate scoring of the writing a challenging task. The reliability rates across each of the 22 writing tasks and grades was 80.6 percent. For the pre-writing, the exact agreement rate was virtually always 95 percent or higher for each category on every task.

7.4 LONG-TERM TREND ASSESSMENTS

7.4.1 Characteristics of the Scoring Guides

Mathematics. The open-ended items in the mathematics bridge assessment were scored on a right/wrong basis.

Reading. The scoring guides for the reading constructed-response questions focused on students' abilities to perform various tasks—for example, identifying the author's message or mood and substantiating their interpretation, making predictions based on given details, and comparing and contrasting information. The guides for the reading items varied somewhat, but typically included a range of scores from unsatisfactory to elaborated. Some of the reading items received secondary scoring based on the reactions or explanations the student provided.

Writing. All of the writing items for the three long-term trend assessments were scored using the primary trait method, which focused on the writer's effectiveness in accomplishing the specific purpose of a given writing task. The primary trait scoring criteria defined five levels of task accomplishment: not rated, unsatisfactory, minimal, adequate, and elaborated. The scoring guide for each item described these levels in detail. In addition, a subset of the items was scored using holistic criteria (see section 7.4.4) and another subset was scored for grammatical correctness (see section 7.4.5).

7.4.2 Training

As with the main assessments, preparation for training readers of the long-term assessment materials entailed reviewing the scoring standards and selecting sample papers. However, several additional considerations were involved in the trend scoring. First, because it was necessary to train the 1992 readers to use the scoring standards from previous assessments (1984 for reading, 1986 for mathematics, 1988 for writing), the same scoring guides and sample papers were used in 1992 training sessions as had been used in the original target years. In addition, scores on booklets from the earlier assessments used in the reliability samples had to be masked so as to prevent readers from being influenced by the previous readers' scores.

During training for the reading and writing assessments, the ETS subject area specialists and NCS team leaders reviewed the scoring guides, elaborated on the rationale underlying the scoring, and discussed illustrative sample responses with the readers. The readers then reviewed scored sample responses and conducted extensive practice scoring.

Because the mathematics items were scored as right, wrong, or omitted, lengthy training of readers was unnecessary. Readers were guided through the scoring rubrics, which listed the correct answer for the items in each of the blocks. Scoring began after a brief episode of practice scoring.

7.4.3 Scoring

After the initial part of the training but before scoring any 1992 long-term trend materials, 25 percent of the 1984 reading and 1988 writing materials were selected for practice scoring. When it was clear that a high degree of interreader agreement had been reached, the readers began scoring the 1992 trend assessment materials.

For the mathematics trend items, readers scored every constructed response in each assessment booklet by gridding the appropriate score in the ovals provided at the bottom of the booklet page. Every tenth booklet was rescored to ensure that the items had been scored accurately. These quality control checks were monitored, and discrepancies were addressed and remediated. Most of the errors found during this check occurred as a result of scores not being gridded.

Two reliability studies were conducted for the reading and writing trend items. For the 1992 material, 25 percent of the constructed responses were scored by a second reader to produce interreader reliability statistics. In addition, to ensure that readers were scoring in keeping with the 1984 and 1988 procedures, 20 percent of the 1984 reading booklets and 25 percent of the 1988 writing booklets were selected and intermixed with the 1992 booklets before the scoring began. Each booklet selected from the 1984 and 1988 assessments had the original scores masked. The readers marked their scores on separate sheets.

Interreader reliability in the reading and writing trend assessments was examined from the two perspectives of comparability to previous assessments and consistency within the assessment itself. To ensure that there was no "drift" in the interpretation of the scoring standards across time, it was important to study the extent to which the scores assigned by the 1992 readers were in agreement with scores assigned by readers in the previous scoring years. For reading, 25 percent of the 1984 papers were randomly selected and given to the 1992 readers to rescore. For writing, 25 percent of the 1988 papers were rescored. The results of these analyses indicated that the level of agreement tended to be fairly high (ranging from 71.2 to 92.0 and averaging 83.5 for writing) and there were no consistent shifts across time in the interpretation of the scoring criteria.

In addition, it was important to know whether the 1992 readers understood and applied scoring standards consistently—in other words, whether they had a high level of agreement with one another. Table 7-2 provides the results of this second set of analyses. In general, the agreement rate on most items was above 80 percent (with the exception of reading at age 13), indicating that an acceptable level of interreader agreement had been achieved.

Table 7-2
Reader Percent Agreement

Assessment	**Number of Unique Items**	**Number of Items by Ranges of Percent Agreement**			
		60-69%	**70-79%**	**80-89%**	**90-100%**
Age 9/Grade 4					
Reading - 1992	6	0	0	6	0
Writing - 1992	5	0	1	3	1
Reading - Trend	6	0	0	4	2
Writing - Trend	5	0	2	2	2
Age 13/Grade 8					
Reading - 1992	7	1	4	2	0
Writing - 1992	6	0	2	4	0
Reading - Trend	7	0	4	3	0
Writing - Trend	6	0	1	5	0
Age 17/Grade 11					
Reading - 1992	8	0	1	6	1
Writing - 1992	6	0	0	5	1
Reading - Trend	8	0	0	6	2
Writing - Trend	6	0	0	6	0

7.4.4 Holistic Scoring of Writing

To gather information about changes across time in the fluency of students' writing, NAEP evaluated some of the 1992 trend writing responses using the holistic scoring method. In this approach, readers evaluated students' writing for general fluency, rather than focusing on specific aspects of writing achievement.

In the NAEP holistic scoring session, two writing trend items were scored from each of the three grade levels (at grade 4, "Flashlight" and "Spaceship"; and at grades 8 and 11, "Food on the Frontier" and "Recreation Opportunities"). The responses were drawn from four assessment years: 1984, 1988, 1990, and 1992.

Preparation

Before the holistic scoring began, the ETS writing specialist explained the theoretical underpinnings of holistic scoring to the team leaders and reviewed the structure of the 6-point scoring guides.

The group reviewed anchor papers chosen in 1988 and used again in 1990 to illustrate each point on the holistic scale. Then, for each item, the ETS writing specialist and tea, leaders read 50 responses drawn from across the four assessment years and assigned scores to these papers. These selected papers were then collated and photocopied for use in training readers.

Training

As the training began, the writing specialist led the readers through a review of the prompt, the accompanying scoring guide, and the six anchor papers. The entire group of readers and the individual teams then discussed the anchor papers at greater length. When the readers were clear about the distinctions among the various score points, the writing specialist chose a small set of responses from the training samples for the readers to score. Practice scoring (and subsequent discussion) continued until the writing specialist was satisfied that the readers had reached a clear and highly uniform understanding of the scoring guide. To maximize interreader reliability, holistic scoring was conducted one item at a time. The interreader reliability for the holistic scoring, average across the six prompts, was 92 percent.

7.4.5 Writing Mechanics Scoring

Mechanics scoring focused on the extent to which the writer demonstrated mastery of the conventions of written English, specifically grammar, spelling, capitalization and punctuation. In addition, the scoring procedures included identifying sentence structures employed and errors in word choice made by the writer.

Preparation

A subset of the responses from the 1992 assessment were selected for mechanics scoring. Every third booklet of the grade-eligible booklets was selected. Black students were oversampled to provide a more stable measure of their performance. In all, 1,808 responses were scored from the 1992 assessment. In addition, 10 percent of the papers that were previously scored for mechanics from the 1988 and 1990 assessments were rescored for reliability. This sample was selected by locating specific booklets from a list generated by ETS and resulted in a rescore of 300 papers for both years.

Before training the mechanics readers, the ETS subject area specialist met with the three team leaders to discuss the scoring guide, review papers that were scored in 1988 and 1990, and score sample responses. After each participant independently scored a set of papers, the group reviewed the individual scores and compared them to the scores assigned in 1988 and 1990. Discrepancies were discussed and resolved. The group then chose a subset of the scored papers to be included in a training packet for the entire group of readers.

Training

Training began with a detailed review of the scoring guide, which was organized into four sections: type of sentence construction, faulty sentence construction, punctuation, and word level categorization. Excerpts from reference sources were distributed and other guides mentioned. After discussing the guide, the group reviewed the training papers.

A subset of 10 percent of the 1988/90 essays scored for mechanics in 1990 but not included in the 1992 rescore was used to train the readers. Copies were made of these training essays to be used for practice. Each reader then individually scored a selected group of essays. The scores were compared among the group, discussions were held when discrepancies occurred, and again references were made to resource materials or to the scored sample papers. When the group was comfortable with the decisions being made, the actual scoring began.

Scoring

In scoring, the mechanics readers marked each paper with a series of symbols, which designated each word or punctuation mark in error and indicated sentence type or faulty sentence construction. Each paper was scored independently by two different readers.

Resolution and quality control were conducted by a team leader who compared the scores marked on separate copies of the responses and resolved any discrepancies. Feedback was provided to each reader and follow-up training discussions were held as necessary. Resolved packets were sent to the NCS word processing department where the text of the papers, along with the assigned scores and identification information, were entered into a computer according to carefully defined specifications. The scoring group proofread the data entry work against the scored papers and any discrepancies were resolved.

7.4.6 Writing Portfolio Scoring

The 1992 NAEP assessment also included a special portfolio study of student writing at the fourth and eighth grades. Students selected for this project submitted, with the assistance of their classroom instructor(s), three samples of their writing from classroom assignments. Each student's teacher was given a teacher questionnaire to complete that asked questions regarding the amount of time the student spent on each assignment, the type of assignment given, and any associated instructional activities used. Students were asked to include a written statement explaining why they selected the writing samples submitted. All these materials were then placed in portfolio folders and sent to NCS by the Westat Supervisor.

The objective in examining the portfolios of students' writing was to learn more about how and what students write under normal classroom conditions. To this end, two types of scoring were conducted. First, each paper included in a portfolio was described according to nine categories relating to writing were examined and coded for each writing sample submitted. These categories were: type of writing process used, time spent in writing the sample, the audience for which the sample was written, whether the topic was chosen by the student, whether the sample was a draft or a final version, whether a computer/typewriter was used, the type or style of writing used, the estimated word count, and an evaluative score. Then, those papers that were classified as narrative, informative, or persuasive were scored. A six-point modified primary trait scoring guide for each of the three genres, developed through pilot testing, were used to evaluate the students' portfolio entries.

Preparation

ETS writing specialists, together with the two team leaders, discussed the descriptive and evaluative scoring guides, reviewed the anchor papers, and aggregated comments from the other reviewers in order to revise the guide. In addition, they choose sample papers to illustrate the six-point scale for each of the three major genres to be scored—narrative, informative, and persuasive.

Training and Scoring

Focused training was preceded by a one-day introduction to trends in current writing instruction. The ETS writing specialists first presented information about writing process methods of instruction, including a videotape of students engaged in various writing activities, to enable the readers to understand the categories used in the descriptive scoring guide. To help them better understand the criteria in the evaluative scoring guide, the readers then received instruction in story theories, grammar, informative writing patterns, and persuasive discourse. The ETS writing specialist then presented a definition of the genre in question, along with an explanation of the scoring guide. The team leaders then led the scorers through a discussion of the guide and review of the anchor papers. When the scorers were clear about the distinction among the various levels in the guide, the leaders gave the scorers a small set of responses from the training samples to score. Practice scoring and subsequent discussion continued until the team leaders and training coordinator were satisfied that the readers had reached a clear understanding of the scoring guide.

Portfolio papers were then scored, with the team leaders resolving any problems that arose. Twenty percent of the portfolios received a blind second scoring, so that interreader reliability could be estimated. Twice a week the team leaders compiled calibration packets, using 10 to 15 papers from the original training packets. Each reader then read and scored these papers. This was done to prevent readers from "drifting" from the original criteria. This procedure continued until scoring was completed. The scores were then recorded on a tally sheet for entry into a computerized data analysis program.

Chapter 8

CREATION OF THE DATABASE, QUALITY CONTROL OF DATA ENTRY, AND THE 1992 NAEP DATABASE PRODUCTS

Alfred M. Rogers, David S. Freund, and John J. Ferris

Educational Testing Service

The data transcription processing of the 1992 NAEP data was conducted by the ETS subcontractor, National Computer Systems (NCS). The processes conducted by NCS, described in Chapter 6, resulted in the transfer to ETS of data files containing response data for students, teachers, and school administrators. At the same time, NCS transferred to ETS subcontractor Westat, Inc., the demographic data needed to derive sampling weights. Westat then provided ETS with data files that included sampling weights for assessed students, excluded students, and schools, principal questionnaire data, school characteristic data, and community characteristic data. Before any analysis could begin, these data files had to be merged into a comprehensive, integrated database. The creation of the database is described in section 8.1.

To evaluate the effectiveness of the quality control of the data entry process, the final database was sampled and verified in detail against the original instruments received from the field. The results of this procedure are given in section 8.2.

The integrated database was the source for the creation of three NAEP database products: the item information database, the restricted-use data files, and the secondary-use data files. These are described in section 8.3.

8.1 CREATION OF THE DATABASE

8.1.1 Merging Files

The transcription process resulted in the generation of 12 sets of data files (long-term trend data and reading, mathematics, and writing cross-sectional data for each of the three student cohorts). Included in each set were student response data, school questionnaire data, and excluded student questionnaire data. A fourth data file contained teacher questionnaire data for the teachers of fourth-grade students assessed in reading or mathematics and teachers of eighth-grade students assessed in mathematics or writing. The process of deriving sampling weights produced an additional three files (assessed student weights, excluded student weights, and school weights) for each of the 12 sets. Before data analyses could be performed, these files had to be integrated into a coherent and comprehensive database.

The database ultimately comprised six files per cohort: four assessed student files (reading, mathematics, writing, and long-term trend), a single school file, and a single excluded student file. The student files were separated by subject area to improve maintenance and efficiency of the databases and data analyses. The student data were created in several steps, merging the student response data with the demographic and community variables, the student weights, and key derived variables such as parental education and home environment composites. Also, teacher data were appended to student records in cases where the student's teacher responded to a teacher questionnaire. In all steps, the booklet serial number was used as the matching criterion.

The school file was created by matching and merging the school questionnaire file with the school weights file and with a file of school variables supplied by Westat, which included demographic information about the school and the community. Some of these data were collected by Westat as part of the principal's questionnaire. The PSU and school code were used as the matching criteria. Each record of the resulting file was formed by merging the weight information with the response data and the school demographic and community data. Since not all schools returned their questionnaires, some of the output records contained only school identifying information and weight information. The school file could be accessed on its own or it could be linked to the assessed and excluded student files through the PSU and school codes.

The excluded student file was the result of merging the excluded student questionnaire file with the excluded student weights file. The booklet serial number was used as the matching criterion.

To ensure that the data were transferred accurately from NCS to ETS and that the processing control parameters used by analysis programs at ETS were properly in place, several quality control procedures were implemented. The principal procedure included matching independently generated frequency distributions computed at NCS with those produced at ETS, using ETS control parameters to process the data. Distributions for all student response questions (approximately 1,500 items) were verified to match.

When the appropriate files had been merged and proper quality control procedures had been completed, the database was ready for analysis. Any time that new data values, such as plausible values, were derived external to the database, they were added to the relevant files using the matching procedures described above. The secondary-use data files were later generated from this database.

8.1.2 Creating the Master Catalog

A critical part of any database is its processing control and descriptive information. A central repository of this information may be accessed by all analysis and reporting programs to provide correct parameters for processing the data fields as well as to provide consistent labeling to identify the results of the analyses. The NAEP master catalog file was designed and constructed to serve both of these purposes.

Each record of the master catalog contains the processing, labeling, classification, and location information for each data field in the NAEP database. The control parameters are used by the access routines in the analysis programs to define the manner in which the data values are to be transformed and processed.

Each data field has a 50-character label in the master catalog describing the contents of and, where applicable, the source of the field. The data fields with discrete or categorical values (e.g., multiple-choice items and constructed-response items, but not weight fields) have additional label fields in the catalog containing 8- and 20-character labels for those values.

The classification area of the catalog record contains distinct fields corresponding to predefined classification categories for the data fields. For a given classification field, a nonblank value indicates the code within that classification category for the data field. This permits the collection of identically classified items or data fields by performing a selection process on one or more classification fields in the catalog.

According to the NAEP design, it is possible for item data fields to appear in more than one student sample and in more than one block within each sample. The location fields of the catalog record contain the age, block and, where applicable, the sequence within the block for each appearance of the data field. (Fields such as plausible values and weights do not contain sequence numbers since these fields are not pertinent to a given block.)

The master catalog file was constructed in parallel with the collection and transcription of the assessment data to be ready for use by analysis programs when the database was created. As new data fields were derived and added to the database, their corresponding descriptive and control information were entered into the catalog.

One of the most important uses of the master catalog was the control of the creation of the secondary-use data files, codebooks, and file layouts. A synopsis of this process is presented in section 8.3.

8.2 QUALITY CONTROL OF NAEP DATA ENTRY FOR 1992

This section describes the evaluation of the data entry process for the 1992 National Assessment. As in past years, this NAEP database was found to be more than accurate enough to support the analyses that were done. Overall, the observed error rates were comparable to those of past assessments; they ranged from 3 errors per 10,000 responses for the student data to 22 errors per 10,000 responses for the excluded student questionnaire data.

The purpose of the analysis reported in this section is to assess the quality of the data resulting from the complete data entry system, beginning with the actual instruments collected in the field and ending with the final machine-readable database used in the analyses. The process involved the selection of instruments at random from among those returned from the field and the comparison of these instruments, character by character, with their representations in the final database. In this way, we were able to measure the error rates in the data as well as the success of the data entry system.

Of course the observed error rate cannot be taken at face value. For example, the sample of school characteristics and policies questionnaires that happened to be selected for close inspection contained no errors at all. To conclude that the entire school characteristics questionnaire database is therefore error-free would be an act of extreme optimism; we may simply have been lucky with this particular random sample. What is needed is an indication of how bad the true error rate might be, given what we observed. Such an indication is provided by confidence limits. Confidence limits indicate how likely it is that a value will fall outside a specified range of values in a specified context or distribution. In our analysis, the specified range is an error rate between zero and some maximum value beyond which we are confident that the true error rate does not lie; the specified context or distribution turns out to be the cumulative binomial probability distribution. An example should demonstrate this technique:

> Let us say that 1,000 booklets were processed, each with 100 characters of data transcribed for a total of 100,000 characters. Let us say further that five of these characters were discovered to be in error in a random sample of 50 booklets that were completely checked; in other words, five errors were found in a sample of 5,000 characters. The following expression may be used to establish the probability that the true error rate is .0025 or less, rather than the single-value estimate of the observed rate of one in a thousand (.001):
>
> $$\sum_{j=0}^{5} \binom{5000}{j} \times .0025^{j} \times (1 - .0025)^{(5000-j)} = .0147$$
>
> This is the sum of the probability of finding five errors plus the probability of finding four errors plus. . . etc. . . plus the probability of finding zero errors in a sample of 5,000 with a true error rate of .0025; that is, the probability of finding five or fewer errors by chance when the true error rate is .0025. Notice that we did not use the size of the database in this expression. Actually, the assumption here is that our sample of 5,000 was drawn from a database that is infinite. The smaller the actual database is, the more confidence we can have in the observed error rate; for example, had there been only 5,000 in the total database, our sample would have included all the data, and the observed error rate would have been the true error rate. The result of the above computation allows us to say, conservatively, that .0025 is an upper limit on the true error rate with 98.53 percent (i.e., 1 - .0147) confidence; that is, we can be quite sure that our true error rate is no larger than .0025.

Virtually all of the data collected for this assessment was machine scanned. The only exception was a set of six booklets used for the long-term trend reading and writing assessments; the format of these booklets was kept the same for comparability with earlier assessments, so these booklets had to be key-entered. As it happened, no errors at all were found in the sample of key-entered booklets selected for quality control.

The individual instruments are briefly discussed in the following sections and a summary table (Table 8-1) gives the upper 99.8 percent confidence limits for the error rates for each of the instruments as well as the sampling information. The confidence limit of 99.8 percent, and

Table 8-1

Summary of Quality Control Error Analysis for NAEP 1992 Data Entry

Instrument/Subsample	Entry Type	Different Booklets	Number of Booklets Sampled	Number of Characters Sampled	Number of Errors	Observed Error Rate	Upper 99.8% Confidence Limit
Student Booklets - Main	Scanned	198	479	39,668	10	.0003	.0006
Student Booklets - Long-term Trend	Keyed	18	43	4,263	0	zero	.0014
	Scanned	8	44	4,455	1	.0002	.0019
Excluded Student Questionnaire	Scanned	1	61	4,941	11	.0022	.0049
Teacher Questionnaire	Scanned	3	35	3,710	6	.0016	.0046
School Characteristics and Policies Questionnaire	Scanned	1	31	3,007	0	zero	.0021

the selection rates noted, were chosen to make these results comparable to those of previous administrations when the same parameters were used.

8.2.1 Student Data

To simplify the selection of student booklets for examination, a method was developed that involved selecting all occurrences of a specified booklet in a randomly selected "stack." A stack is a unit of collection containing anywhere from 10 to 100 booklets, but typically between 50 and 60 booklets, in an assortment related to the spiraling technique used to distribute the booklets. The selection method was designed to yield approximately the same number of each booklet but, due to the variability in the size and contents of the stacks, there was somewhat more variation in the numbers of booklets selected than in previous assessments. However, all of the booklets were sampled in adequate numbers and the average rate of selection was 1/270, which is higher than the 1/400 used in past assessments. The few errors found during this quality control examination did not cluster by booklet number, so there is no reason to believe that the variation in numbers of booklets selected had a significant effect on the estimates of overall error rate confidence limits reported below.

About 150,000 students were assessed across all samples in this assessment. The error rates were consistently low in all subject areas and across all three grades. The overall quality of the data was very high.

8.2.2 Excluded Student Questionnaire Data

A total of 11,847 excluded student questionnaires was scanned in this assessment and were sampled for quality control at a rate of about one in 200 questionnaires. These questionnaires and the teacher questionnaires had nearly ten times the error rate of the student data. It should be noted that this difference reflects the extremely low error rates in the student data rather than an alarming error rate in the questionnaire data; the questionnaire data were more than accurate enough to support the analyses. While it is possible that something in the design of these instruments contributed to the higher error rate, it seems more likely that they were not completed with the same care that the students exercised in working with their booklets. The great majority of the problems came from erasures that were not correctly diagnosed by the scanner.

8.2.3 Teacher Questionnaire Data

There were 3,655 teacher questionnaires collected and scanned in this national assessment. One percent of these questionnaires were sampled for the quality control procedure. As mentioned in the paragraph above, the error rates were higher than those of the student data, although comparable to those found in past assessments. The errors found involved misreadings of erasures and multiple responses by the scanner.

8.2.4 School Characteristics and Policies Questionnaire Data

There were 2,357 school characteristics and policies questionnaires collected in this assessment. They were sampled at a rate of about one in 75. No errors were found in the questionnaires that were checked.

8.3 NAEP DATABASE PRODUCTS

The NAEP database described to this point serves primarily to support analysis and reporting activities that are directly related to the NAEP contract. This database has a singular structure and access methodology that is integrated with the NAEP analysis and reporting programs. One of the directives of the NAEP contract is to provide secondary researchers with a nonproprietary version of the database that is portable to any computer system. In the event of transfer of NAEP to another client, the contract further requires ETS to provide a full copy of the internal database in a format that may be installed on a different computer system.

In fulfillment of these requirements, ETS provides three sets of database products: the item information database, the restricted-use data files, and the secondary-use data files. The contents, format and usage of these products are documented in the publications listed under the appropriate sections below.

8.3.1 The Item Information Database

The NAEP item information database contains all of the descriptive, processing, and usage information for every assessment item developed and used for NAEP since 1970. The primary unit of this database is the item. Each NAEP item is associated with different levels of information, including usage across years and age cohorts, subject area classifications, response category descriptors, and locations of response data on secondary-use data files.

The item information database is used for a variety of essential NAEP tasks: providing statistical information to aid in test construction, determining the usage of items across assessment years and ages for trend and cross-sectional analyses, labeling summary analyses and reports, and organizing items by subject area classifications for scaling analysis.

The creation, structure, and use of the NAEP item information database for all items used up to and including the 1992 assessment are fully documented in the NAEP publications, *A Guide to the NAEP Item Information Database* (Rogers, Barone, & Kline, 1990) and *A Primer for the NAEP Item Information Database* (Rogers, Kline, Barone, Mychajlowycz, & Forer, 1989).

The procedures used to create the 1992 version of the item information database are the same as those documented in the guide. The updated version of the guide also contains the subject area classification categories for the cognitive items.

8.3.2 The Restricted-use Data Files

The restricted-use data files are for the internal use of the NAEP contractor. They contain a complete copy of the internal NAEP respondent database in a structured, documented, and portable format.

The internal database is maintained in a compressed format to conserve computing resources and to increase analysis efficiency. The access methods developed for this database locate data fields dynamically during the execution of analysis programs. The restricted-use data files, on the other hand, are "rectangular" in structure; each data field is in the same location on every record within a file. This static data definition, while not efficient from a computing resource standpoint, is much easier to document and is not dependent on any computing machinery, operating system, or data access method.

The restricted-use data files serve several critical purposes. They provide an archive for all respondent data collected and derived for NAEP since 1970. They ensure compatibility of usage by expressing the data in consistent, rectangular formats. Their portability greatly facilitates transition of the respondent database to future NAEP contractors. The accompanying data file layouts and codebooks provide a standardized, comprehensive reference source for NAEP staff.

The contents and formats of the NAEP restricted-use data files are documented in the NAEP publication *A Guide to the NAEP Restricted-use Data Files* (Rogers, Barone, & Kline, 1989).

The procedures used to create the restricted-use data files for the 1992 assessment are the same as those used to create the secondary-use data files. Since the secondary-use data file distribution package contains more products, the generation procedures will be described in the following section.

8.3.3 The Secondary-use Data Files

The secondary-use data files are designed to enable any researcher with an interest in the NAEP database to perform secondary analysis on the same data as those used at ETS. They differ from the restricted-use data files in one important respect: all subregional identification information has been encrypted or excluded in order to maintain the confidentiality of the states, schools, and students who participated in the assessment.

The three elements of the distribution package are the data files, the printed documentation, and the microfiche copies of the assessment instruments. A set of files for each sample or instrument contains the response data file, a file of control statements that will generate an SPSS-X system file, a file of control statements that will generate a SAS system file, and a machine-readable catalog file containing control and descriptive information, intended for the user who does not use either SAS or SPSS-X. The printed documentation consists of four volumes: a guide to the use of the data files, and a set of data file layouts and codebooks for each of the three age cohorts (see *The NAEP 1992 Secondary-use Data Files User Guide* [Rogers, Kline, Johnson, Mislevy, & Rust, 1994]).

The remainder of this section summarizes the procedures used in generating the data files and related materials.

File Definition

The design of the 1992 assessment perpetuates two features of the 1990 assessment design: the focused- BIB booklet design and the direct matching of teacher questionnaires to student assessment instruments.

The focused-BIB design within the main assessment isolates the primary subject areas to separate groups of booklets. This permits the division of the main sample into subject-specific subsamples. The data files generated from these subsamples need only contain the data that is relevant to their corresponding subject areas and are therefore smaller and more manageable than their counterparts in previous assessments.

According to the design of the 1984, 1986, and 1988 assessments, only a sample of the teachers of the assessed students were asked to fill out the teacher questionnaires. The large size of the secondary-use main student files and the relatively low matching rate between students and teachers made it impractical if not physically prohibitive to produce a complete file with student and teacher information. Both the 1984 and 1986 secondary-use data packages had separate teacher data files which could be linked to the student data files for analysis. The teacher file in the 1988 secondary-use data package contained not only the teacher response data, but also the data from the students who could be matched to teacher questionnaires. This type of file was more appropriate for the analysis of teacher data because it defined the student as the unit of observation.

The intent of the 1992 assessment design was to collect data from the reading, mathematics, or writing teachers of the main assessment students at specified age/grade levels who were administered reading, mathematics, or writing booklets. A portion of the teacher questionnaire contained questions that were directly related to each matched student. This change in the design afforded a very high matching rate between student and teacher data. Therefore, for those subject areas in each age/grade cohort for which teacher data were collected, the teacher responses were appended to each student record in the secondary-use data files.

Definition of the Variables

The selection and arrangement of data fields, or variables, in each file was the next issue addressed. The initial step in this process was the generation of a LABELS file of descriptors of the variables for each data file to be created. Each record in a LABELS file contains, for a single data field, the variable name, a short description of the variable, and processing control information to be used by later steps in the data generation process. This file could be edited for deletion of variables, modification of control parameters, or reordering of the variables within the file. The LABELS file is an intermediate file only; it is not included on the released data files.

The first program in the processing stream, GENLYT, produced a printed layout for each file from the information in its corresponding LABELS file. These layouts were initially reviewed for the selection and ordering of the variables. Variables were excluded from secondary-use data files if they were classified as either confidential or nonapplicable.

The confidential variables included any descriptor or code that could be used to identify individual states, schools, or students in the NAEP sample. The PSU, school, teacher, and student identification codes used internally by ETS and Westat were "scrambled" according to specific algorithms to obtain new codes for use in linking the files together. These new codes were put on the secondary-use files in lieu of the original codes.

The nonapplicable variables were found mostly in the student database. In the database used for analysis and reporting, the long-term trend samples were combined with the main sample. Therefore, many of the variables that applied to the main sample students did not apply to the trend sample students, and vice versa. Similarly, within the main assessment sample, students who were administered booklets in one subject area had no derived variables pertaining to the other subject areas. When the data for these samples were separated into different datasets for the secondary-use data files, these nonapplicable variables were excluded.

The variables on all data files were grouped and arranged in the following order: identification information, weights, derived variables, and response data. On the student data files, these fields were followed by the proficiency scale scores and teacher response data, where applicable. The identification information is taken from the front covers of the instruments. The weight data include sample descriptors, selection probabilities, and replicate weights for the estimation of sampling error. The derived data include sample descriptions from other sources and variables that are derived from the response data for use in analysis or reporting.

For each subject area of the main assessment, the item response data within each block were left in their order of presentation. The blocks, however, were arranged according to the following scheme: common background, subject-related background, and cognitive blocks in ascending numerical order. The responses to cognitive blocks that were not present in a given booklet were left blank, signifying a condition of "missing by design."

In order to process and analyze the spiral sample data effectively, the user must also be able to determine, from a given booklet record, which blocks of item response data were present and their relative order in the instrument. This problem was remedied by the creation of a set of control variables, one for each block, which indicated not only the presence or absence of the block but its order in the instrument. These control variables were included with the derived variables.

Data Definition

To enable the data files to be processed on any computer system using any procedural or programming language, it was desirable that the data be expressed in numeric format. This was possible, but not without the adoption of certain conventions for reexpressing the data values.

As mentioned in section 8.1, the responses to all multiple-choice items were transcribed and stored in the database using the letter codes printed in the instruments. This scheme afforded the advantage of saving storage space for items with 10 or more response options, but at the expense of translating these codes into their numeric equivalents for analysis purposes. The response data fields for most of these items would require a simple alphabetic-to-numeric conversion. However, the data fields for items with 10 or more response choices would require "expansion" before the conversion, since the numeric value would require two column positions. One of the processing control parameters on the LABELS file indicates whether or not the data field is to be expanded before conversion and output.

The ETS database contained special codes to indicate certain response conditions: "I don't know" responses, multiple responses, omitted responses, not-reached responses, and unresolvable responses, which included out-of-range responses and responses that were missing due to errors in printing or processing. The primary trait scores for the reading essay and writing items included additional special codes for ratings of "illegible," "off task," and nonrateable by the scorers. All of these codes had to be reexpressed in a consistent numeric format.

The following convention was adopted and used in the designation of these codes: The "I don't know" and nonrateable response codes were always converted to 7; the omitted response codes were converted to 8; the "not-reached" response codes were converted to 9; the multiple response codes were converted to 0; the "illegible" codes were converted to 5; and the "off task" codes were converted to 6. The out-of-range and missing responses were coded as blank fields, corresponding to the "missing by design" designation.

This coding scheme created conflicts for those multiple-choice items that had seven or more valid response options as well as the "I don't know" response and for those open-ended items whose primary trait scoring guide had five or more categories. These data fields were also expanded to accommodate the valid response values and the special codes. In these cases, the special codes were "extended" to fill the output data field: The "I don't know" and nonrateable codes were extended from 7 to 77, omitted response codes from 8 to 88, etc.

Each numeric variable on the secondary-use files was classified as either continuous or discrete. The continuous variables include the weights, proficiency values, identification codes, and item responses where counts or percentages were requested. The discrete variables include those items for which each numeric value corresponds to a response category. The designation of "discrete" also includes those derived variables to which numeric classification categories have been assigned. The open-ended items were treated as a special subset of the discrete variables and were assigned to a separate category to facilitate their identification in the documentation.

Data File Layouts

The data file layouts, as mentioned above, were the first user product to be generated in the secondary-use data files process. The generation program, GENLYT, used a LABELS file as input and produced a printable file. The LAYOUT file is little more than a formatted listing of the LABELS file.

Each line of the LAYOUT file contains the following information for a single data field: sequence number, field name, output column position, field width, number of decimal places, data type, value range, key or correct response value, and a short description of the field. The sequence number of each field is implied from its order on the LABELS file. The field name is an 8-character label for the field that is to be used consistently by all secondary-use data files materials to refer to that field on that file. The output column position is the relative location of the beginning of that field on each record for that file, using bytes or characters as the unit of measure. The field width indicates the number of columns used in representing the data values for a field. If the field contains continuous numeric data, the value under the number of decimal places entry indicates how many places to shift the decimal point before processing data values.

The data type category uses three codes to designate the nature of the data in the field: Continuous numeric data are coded "C"; discrete numeric data are coded "D"; open-ended item data are coded "O." Additionally, the discrete numeric fields that include "I don't know" response codes are coded "DI," the short constructed-response items that include nonrateable response codes are coded "OI," and the extended constructed-response items are coded "OE." If the field type is discrete numeric, the value range is listed as the minimum and maximum permitted values separated by a hyphen to indicate range. If the field is a response to a scorable item, the correct option value, or key, is printed. A range of correct options was indicated for those professionally scored items that were treated with cutoff scoring for IRT scaling. Finally, each variable was further identified by a 50-character descriptor.

Data File Catalogs

The LABELS file contains sufficient descriptive information for generating a brief layout of the data file. However, to generate a complete codebook document, substantially more information about the data is required. The CATALOG file provides most of this information.

The CATALOG file is created by the GENCAT program from the LABELS file and the 1992 master catalog file. Each record on the LABELS file generates a CATALOG record by first retrieving the master catalog record corresponding to the field name. The master catalog record contains usage, classification, and response code information, prefixed by the positional information from the LABELS file: field sequence number, output column position, and field width. Like the LABELS file, the CATALOG file is an intermediate file and is not included on the release data files.

The information for the response codes, also referred to as "foils," consists of the valid data values for the discrete numeric fields, and a 20-character description of each. The GENCAT program uses additional control information from the LABELS file to determine if extra foils should be generated and saved with each CATALOG record. The first flag controls generation of the "I don't know" or nonrateable foil; the second flag regulates omitted or not-reached foil generation; and the third flag denotes the possibility of multiple responses for that field and sets up an appropriate foil. All of these control parameters, including the expansion flag, may be altered in the LABELS file by use of a text editor, in order to control the generation of data or descriptive information for any given field.

The LABELS file supplies control information for many of the subsequent secondary-use data processing steps. The CATALOG file provides detailed information for those and other steps.

Data Codebooks

The data codebook is a printed document containing complete descriptive information for each data field. Most of this information originates from the CATALOG file; the remaining data came from the COUNTS file and the IRT parameters file.

Each data field receives at least one line of descriptive information in the codebook. If the data type is continuous numeric, no more information is given. If the variable is discrete numeric, the codebook lists the foil codes, foil labels, and frequencies of each value in the data file. Additionally, if the field represents an item used in IRT scaling, the codebook lists the parameters used by the scaling program.

Certain blocks of cognitive items in the 1992 assessment that are to be used again in later assessments for trend comparisons have been designated as nonreleased. In order to maintain their complete confidentiality, the descriptions of these items and their response categories have been substituted with generic labels in the file layouts, data codebooks, and user guide.

The frequency counts are not available on the catalog file, but must be generated from the data. The GENFREQ program creates the COUNTS file using the field name to locate the variable in the database, and the foil values to validate the range of data values for each field. This program also serves as a check on the completeness of the foils in the CATALOG file, as it flags any data values not represented by a foil value and label.

The IRT parameter file is linked to the CATALOG file through the field name. Printing of the IRT parameters is governed by a control flag in the classification section of the CATALOG record.

The LAYOUT and CODEBOOK files are written by their respective generation programs to print-image disk data files. Draft copies are printed and distributed for review before the production copy is generated. The production copy is printed on an IBM 3800 printer that uses laser-imaging technology to produce high-quality, reproducible documentation.

Control Statement Files for Statistical Packages

An additional requirement of the NAEP contract is to provide, for each secondary-use data file, a file of control statements each for the SAS and SPSS-X statistical systems that will convert the raw data file into the system data file for that package. Two separate programs, GENSAS and GENSPX, generate these control files using the CATALOG file as input.

Each of the control files contains separate sections for variable definition, variable labeling, missing value declaration, value labeling, and creation of scored variables from the cognitive items. The variable definition section describes the locations of the fields, by name, in the file, and, if applicable, the number of decimal places or type of data. The variable label identifies each field with a 50-character description. The missing value section identifies values of those variables that are to be treated as missing and excluded from analyses. The value labels correspond to the foils in the CATALOG file. The code values and their descriptors are listed for each discrete numeric variable. The scoring section is provided to permit the user to generate item score variables in addition to the item response variables.

Each of the code generation programs combines three steps into one complex procedure. As each CATALOG file record is read, it is broken into several component records according to the information to be used in each of the resultant sections. These record fragments are tagged with the field sequence number and a section sequence code. They are then sorted by section code and sequence number. Finally, the reorganized information is output in a structured format dictated by the syntax of the processing language.

The generation of the system files accomplishes the testing of these control statement files. The system files are saved for use in special analyses by NAEP staff. These control statement files are included on the distributed data files to permit users with access to SAS and/or SPSS-X to create their own system files.

Machine-readable Catalog Files

For those NAEP data users who have neither SAS nor SPSS-X capabilities, yet require processing control information in a computer-readable format, the distribution files also contain machine-readable catalog files. Each machine-readable catalog record contains processing control information, IRT parameters, and foil codes and labels.

PART II

The Analysis of 1992 NAEP Data

Chapter 9

OVERVIEW OF PART II: THE ANALYSIS OF 1992 NAEP DATA[1]

James E. Carlson and Spencer S. Swinton

Educational Testing Service

This chapter presents an overview of the analyses conducted on the 1992 NAEP data, focusing on the common elements of the analyses used across the subject areas of the assessment. Details by subject area are provided in Chapters 12 through 15.

As described in Chapter 1, the 1992 NAEP comprised three major components as well as several special studies. One component encompassed major assessments in reading, mathematics, and writing, providing detailed information about student proficiency at the fourth- eighth- and twelfth-grade levels of private and public schools. Long-term trend assessments of reading, mathematics, science, and writing at the same three levels constituted a second component. The third major component was the Trial State Assessment at the fourth-grade level in reading and mathematics and at the eighth-grade level in mathematics. Technical details of the Trial State Assessment are not included in this report but are presented in two other publications: the *Technical Report of the NAEP 1992 Trial State Assessment Program in Mathematics* (Johnson, Mazzeo, & Kline, 1993), and the *Technical Report of the NAEP 1992 Trial State Assessment Program in Reading* (Johnson, Mazzeo, & Kline, 1994).

Special studies in the 1992 National Assessment included:

- an oral reading assessment (the Integrated Reading Performance Record study);
- in mathematics, assessment of students' ability to perform tasks involving estimation of quantities;
- a targeted assessment of complex problem-solving in pre-algebra and algebra;
- a fourth-grade mathematics study designed to measure the effect of changing the type of calculator used; and
- an assessment of students' school-based writing (the Writing Portfolio study).

Descriptions of the analyses carried out in these special studies are presented in the subject area chapters.

---

[1] Nancy Allen, Eugene Johnson, John Mazzeo, and Rebecca Zwick contributed to this chapter.

One significant innovation in the analyses of NAEP for 1992 as compared to previous assessments is the use of item response theory (IRT) methodology in the scaling of the writing assessment. This methodology, which is new to the writing assessment, resulted in the implementation of new scales in writing, both for long-term trend and the main assessment. For the writing trend assessment, this involved IRT scaling of data from the last four writing assessments.

Results from the analyses described in the following chapters were reported in the *NAEP 1992 Reading Report Card for the Nation and the States* (Mullis, Campbell, & Farstrup, 1993); the *NAEP 1992 Mathematics Report Card for the Nation and the States* (Mullis, Dossey, Owen, & Phillips, 1993); the *NAEP 1992 Writing Report Card* (Applebee, Langer, Mullis, Latham, & Gentile, 1994); *NAEP 1992 Trends in Academic Progress: Achievement of U.S. Students in Science, 1969 to 1992; Mathematics, 1973 to 1992; Reading, 1971 to 1992; Writing, 1984 to 1992* (Mullis, Dossey, Campbell, Gentile, O'Sullivan, & Latham, 1994); and *Windows into the Classroom: NAEP's 1992 Writing Portfolio Study* (Gentile & Martin-Rehrmann, 1994).

9.1 SAMPLES OF STUDENTS

Because the samples of students included in the 1992 NAEP assessment are listed and described in detail in Chapter 1, only a brief description of these samples is given here. The 1992 national samples were of three general types: main NAEP samples, which were based on a common set of assessment procedures, including calendar-year age definitions; long-term trend samples, the purpose of which was to provide links to earlier assessments; and special study samples, which were from the same populations as the main NAEP samples but used different administration procedures, specifically designed for the special studies.

To shorten the timetable for reporting results, the period for national data collection was shortened in 1992 from the five-month period (January through May) used in 1990 to a three-month period in the winter (January through March, corresponding to the period used for the winter half-sample of the 1990 National Assessment). To enhance the coverage of the subject areas assessed, the number of cognitive items administered in the assessment was increased for 1992. In addition, the 1992 mathematics assessment contained more constructed-response items and more items requiring extended constructed responses, as opposed to short-answer responses.

The 1992 analyses of long-term trend data extended the trend lines commencing in 1971 in reading, 1973 in mathematics, 1970 in science, and 1984 in writing.

As described in Chapters 1 and 4, for each subject area, blocks of items were used to create a large number of different assessment booklets according to a focused balanced incomplete block (focused-BIB) design. The focused-BIB design provided for booklets that typically included three blocks of cognitive items in a single subject area, as well as background items. The blocks of cognitive items for reading and mathematics included both multiple-choice and constructed-response items; cognitive items in the writing blocks were all constructed-response. With some variation, especially with respect to blocks created for special studies, each block of cognitive items typically appeared in the same number of booklets. To balance possible block position main effects, each block appeared an equal number of times in each position. In addition, the BIB design required that each block of items be paired in a booklet with every other block of items exactly once. Some additional booklets were included in the main assessment for the special studies.

9.2 ANALYSIS STEPS

Because the analysis methods are not identical across subject areas, a separate chapter has been included for each area. The procedures used depended on whether test items were scored dichotomously (right versus wrong) or polytomously (more than two categories of response) and whether links across age groups and/or across assessments were required. Basic procedures common to most or all of the subject area analyses are summarized here. The order is essentially that in which the procedures were carried out.

9.2.1 Preparation of Final Sampling Weights

Because NAEP uses a complex sampling design (Chapter 3) in which students in certain subpopulations have different probabilities of inclusion in the sample, the data collected from each student must be assigned a weight to be used in analyses. The 1992 NAEP weights were provided by Westat, Inc., the NAEP subcontractor in charge of sampling. Detailed information about the weighting procedures is available both in Chapter 10 and in a report prepared by Westat personnel (Wallace & Rust, 1994).

9.2.2 Analysis of Item Properties: Background and Cognitive Items

The first step in the analysis of the 1992 data was item-level analysis of all instruments. In the case of mathematics, certain item analysis procedures were performed upon receipt of approximately the first 25 percent of data for purposes of early screening. This "early" item analysis, which used preliminary student weights, helped ensure meeting the early reporting deadline for mathematics. When complete data were received, the full item analysis procedures were applied to mathematics data, using the same procedures as for the other subject areas, and using the final weights. Item analyses were performed separately for each grade level on each item in each subject area. Each block of items was analyzed separately, by grade, with the total score on the block used as the criterion score for statistics requiring such a score.

Background Items

For each NAEP background item, the unweighted and weighted percent of students who gave each response were examined, as well as the percent of students who omitted the item and the percent who did not reach the item. The number of respondents was also tabulated. These preliminary analyses were conducted within age/grade cohort and within major reporting categories. If unexpected results were found, the item data and the encoding of responses were rechecked.

Cognitive Items

All NAEP cognitive items were subjected to analyses of item properties. These analyses included conventional item analyses and analyses of differential item functioning, and incorporated examinee sampling weights. Item analysis was conducted at the block level so that "number-correct" scores for an item are those for the block containing that item. Because of the inclusion of polytomously scored items in the cognitive instruments, it was necessary to develop special procedures for these items. The resulting statistics are analogous to those for the dichotomously scored items, as listed below.

Dichotomously Scored Items. These items were analyzed using standard procedures that result in a report for each item that includes:

- for each option of the item, for examinees omitting and not reaching the item, and for the total sample of examinees:
    - number of examinees,
    - percentage of examinees,
    - mean of number-correct scores, and
    - standard deviation of number-correct scores;
- p+, the proportion of examinees that received a correct score on the item (ratio of number correct to number correct plus wrong plus omitted);
- Δ, the inverse-normally transformed p+ scaled to mean 13 and standard deviation 4;
- the biserial correlation coefficient between the item and the number-correct scores; and
- the point-biserial correlation coefficient between the item and the number-correct scores.

Polytomously Scored Items. Enhanced procedures were employed for polytomously scored items. First, methods parallel to those for dichotomously scored items resulted in values reported for each distinct response category for the item. For example, a constructed-response item with four response categories would initially have six categories (not-reached, omitted, and the four valid response categories). Another set of statistics resulted from mapping the response categories (excluding not-reached) into a new set of categories reflecting the scoring guide for the item. A constructed-response item with ordered categories, for example, would be mapped into a set of integers in a corresponding order. The scoring rubric could result in the collapsing of (combining of) some response categories.

For these items the following statistics, analogous to those for dichotomously scored items, were computed:

- in place of p+, the ratio of the mean item score to the maximum-possible item score was used;
- in place of Δ, the ratio of the mean item score to the maximum-possible item score underwent the same transformation as that used on p+ to get Δ for dichotomously scored items;
- the polyserial correlation coefficient was used in place of the biserial; and
- the Pearson correlation coefficient was used in place of the point-biserial.

Tables of Item-level Results

Tables were created of the percentages of students choosing each of the possible responses to each item within each of the samples administered in 1992. The results for each item were cross-tabulated against the basic reporting variables such as region, gender, race/ethnicity, public/private school, and parental education. All percentages were computed using the sampling weights.

The sampling variability of all population estimates was obtained by the jackknife procedure used by ETS in previous assessments. In addition to an estimator of the sampling variance of these statistics, an indication of the number of degrees of freedom to attribute to the estimated variance was provided for selected key populations. Details of these procedures are presented in Chapter 10.

Differential Item Functioning Analysis of Cognitive Items

Differential item functioning (DIF) analysis refers to procedures to assess whether items are differentially difficult for different groups of examinees. DIF procedures typically control for overall between-group differences on a criterion, usually by matching examinees from the two groups on overall test scores. Between-group performance on each item is then compared within sets of examinees having the same total test scores.

The approach to DIF analysis that has been adopted by ETS is based on the Mantel-Haenszel test (Holland, 1985; Holland & Thayer, 1988). The procedure involves identifying a focal group (e.g., Black, Hispanic, or female students) and a reference group (e.g., White or male students), and conducting an analysis to support inferences as to whether the item is differentially difficult (either harder or easier) in the focal group as compared to the reference group, after matching on a criterion variable. The Mantel-Haenszel procedure is performed separately for each item and the value of a test statistic, the Mantel-Haenszel D-DIF statistic, is computed. In the application of the ETS corporate-wide guidelines for interpretation of the results of the Mantel-Haenszel procedure, the size and statistical significance of the Mantel-Haenszel D-DIF statistics are used to classify items as "A" items, which are considered free of

DIF, "B" items, which show some evidence of DIF, or "C" items, which show notable evidence of DIF (Zwick & Grima, 1991; Zwick & Ercikan, 1989).

Identifying items as differentially functioning for any particular demographic groups using this method gives no indication of the fairness of the items. ETS guidelines require an examination of "C" items by content and sensitivity experts to make judgments about the fairness of these items. In the 1992 NAEP analyses, "C" items were examined by a committee of subject-area and testing experts to determine whether they should be retained for use in the assessments. Should examination of an item by experts indicate that it is problematic that item may be dropped from the subsequent analyses. If all items that were identified as "C" items by the DIF statistics were eliminated from the assessment instruments, information about group differences in performance would be lost.

The ETS DIF procedure has been adapted to account for the special characteristics of NAEP data resulting from student and item sampling. Because research results (Zwick & Grima, 1991) strongly suggest that sampling weights should be used in conducting DIF analyses, the weights were used. The results of the DIF analyses are shown in Appendix L.

The items were analyzed for DIF using an adaptation of the ETS Mantel-Haenszel procedure. Because all examinees do not take all items in NAEP assessments, it was necessary to modify the above-described ETS procedure. The effect of NAEP's complex sampling plan on properties of the Mantel-Haenszel D-DIF statistic has been investigated by Nelson and Zwick (1989) and Allen and Donoghue (1993). The results of this research suggest using the block score, pooled over booklets, as the matching variable. Hence that procedure, described in Allen and Donoghue, was used for DIF analysis of the cognitive items. Because procedures and software for DIF analysis of polytomously scored items using the full information in all response categories were not fully developed at the time of the analyses of the 1992 NAEP assessments, such items were dichotomized and analyzed using the ETS procedure, which assumes dichotomously scored responses.

9.2.3 Scaling

Scales based on item response theory (IRT) were derived for each subject area. A single scale was used for summarizing long-term trends at each grade level in each of the subject areas. One scale was used for the writing main assessment data; five and three scales were created respectively for mathematics and reading main assessment data. NAEP uses the methodology of multiple imputations ("plausible values") to estimate characteristics of the proficiency distributions. Chapter 11 describes in detail the theoretical underpinnings of NAEP's scaling methods and the required estimation procedures. The basic analysis steps are outlined here.

1) Use the NAEP-BILOG/PARSCALE computer program (described in Chapter 11) to estimate the parameters of the item response functions on an arbitrary scale. This program uses an IRT model incorporating the two- and three-parameter logistic forms used in previous assessments for dichotomously scored items and the generalized partial credit form for polytomously scored items. In order to select starting values for the iterative parameter-estimation procedure for each dataset, the program is first run to convergence, imposing the condition of a fixed normal prior

distribution of the proficiency variable. Once these starting values are computed, the main estimation runs are done with the prior "freed." That is, no prior assumption about the shape of the proficiency distribution is made. In analyses involving more than one population, estimates of parameters are made with the overall mean and standard deviation of all subjects' proficiencies specified to be zero and one, respectively. This is referred to in using the program as "centering."

2) Use a version of the MGROUP program (described in Chapter 11), which implements the method of Mislevy (see Chapter 11 or Mislevy, 1991) to estimate proficiency distributions for each student on an arbitrary scale, based on the item parameter estimates and the student's responses to cognitive items and background questions.

3) Determine the appropriate metric for reporting the results and transform the results as needed. This includes the definition of composite scores for mathematics and reading, and the linking of current scales to scales from the past, as well as the selection of the mean and variance of new scales.

4) Use random draws from these proficiency distributions ("plausible values," in NAEP terminology) for computing the statistics of interest, such as means for demographic groups.

5) Use the jackknife procedure to estimate the standard errors of the mean proficiencies for the various demographic groups.

After the NAEP BILOG/PARSCALE item parameter estimation was completed for each scale, careful examination of the results was necessary to assure that all items can be appropriately incorporated into the scale (fit the model). The primary means of accomplishing this is to generate plots of empirical versus theoretical item response curves. The theoretical curves are plots of the response functions based on the estimates of the item parameters. The empirical curves are not model-based. Rather, examinees are divided into categories according to estimated proficiency scores, and the proportion answering the item correctly (or scored in a given score category for polytomously scored items) are plotted at each of the selected proficiency levels (midpoints of the categories). For good-fitting items the empirical and theoretical curves are close together. Hence items for which this is not true are carefully examined before proceeding with the next step in the analysis. Examples of the plots are provided in the subject-area chapters.

As explained in Chapter 11, the plausible values obtained through the IRT approach are not optimal estimates of individual proficiency; instead, they serve as intermediate values to be used in estimating subpopulation characteristics. Under the assumptions of the scaling models, these subpopulation estimates will be consistent, which would not be true of subpopulation estimates obtained by aggregating optimal estimates of individual proficiency.

Scaling the Test Items

The data from both the trend and the main assessment samples were scaled using IRT models. For dichotomously scored items two- and three-parameter logistic forms of the model were used, while for polytomously scored items the general partial credit model form was used. The long-term trend and the main assessment data were scaled separately. As described above, in the course of the scaling, analyses of model fit to detect and correct aberrations related to differential functioning across subpopulations, or item-by-time interactions, were carried out.

Long-term Trend Scaling. Item parameters for mathematics, reading, and science trends were reestimated, separately for each age group using the data from the 1990 assessment as well as data from the 1992 assessment. The resulting scales, based on these reestimated item parameters, were then linked to the existing long-term trend scales.

Scaling with IRT models had not been used in past assessments for writing because software incorporating models for polytomously- and dichotomously scored items was not yet available. Hence item parameters for this subject were estimated using the data from the 1984, 1988, and 1990 assessments as well as that from the 1992 assessment.

Main Assessment Scaling. Each of the main assessments of mathematics, reading, and writing have special characteristics that determine the procedures that were followed for main assessment scaling and that determine the relationship between the resultant scales and previously created scales for the subject area. The key consideration was the degree of similarity between the 1992 assessment of the subject area and earlier assessments in terms of the populations assessed and the characteristics of the assessment instrument used.

The frameworks for the different subject areas dictate differences in the numbers of scales. For mathematics and reading item parameter estimation was performed separately for each of several scales defined in their frameworks, using data from all the age/grade samples for which the scale was defined. Separate calibrations for scales were not necessary for writing because the writing framework defines only one scale. Item parameter estimates on a provisional scale were obtained using the NAEP BILOG/PARSCALE program. The fit of the IRT model to the observed data was examined within each scale by comparing the theoretical item response functions with the theoretical curves, as described above. For each item presented at more than one grade, the empirical item response functions for each grade were compared. Any item for which the empirical item response functions differed significantly by grade was treated as a set of distinct items; that is, a different item for each grade. Items that displayed significant lack of fit across one or more grades were omitted from scaling for those grades.

Generation of Plausible Values for Each Scale

After the scales were developed, plausible values were drawn from the predictive distribution of proficiency values for each student. For the trend and writing scales, the plausible values were computed separately for each age or age/grade group and year, and were based on the student's responses to the items going into the scale as well as on the values of a set of conditioning variables that were important for the reporting of proficiency scores. For the mathematics and reading main scales, vectors of multivariate plausible values were drawn from the joint distribution of scale proficiency values for the assessed student. These multivariate plausible values were computed separately for each age/grade and reflected the dependency between scale proficiencies by utilizing shared variation among the scales. All plausible values were later rescaled to the final proficiency scale metric using appropriate linear transformations.

The variables used in conditioning a given main assessment scale or group of scales included a broad spectrum of background, attitude, and experiential variables and composites of such variables and explicitly included all standard reporting variables. The set of conditioning variables were defined with the aim of holding to low levels secondary biases in analyses involving a broad range of variables not included in the conditioning model. To minimize potential convergence problems for the main assessment scales, the original background variable contrasts were standardized and transformed into a set of linearly independent variables by extracting principal components from the correlation matrix of the original contrast variables. The principal components, rather than the original variables, were used as independent variables in the conditioning model for those scales. Trend scales used the same or similar sets of conditioning variables that were used when the scales were originally constructed. Details of the conditioning process and of the NAEP-BGROUP and NAEP-CGROUP computer programs that implement the process are presented in Chapter 11.

Definition of Composites for the Multivariate Scales

In addition to the plausible values for each scale, composites of the individual mathematics and reading main assessment scales were created as measures of overall proficiency within these two subject areas. These composites were weighted averages of the plausible values of the individual scales. The weights reflected the relative importance of the scales and were provided in the frameworks developed by the subject area committees. The weights are approximately proportional to the number of items in each scale at a given age level.

Transformation to the Reporting Metric

Mathematics short-term trend and mathematics, reading, and science long-term trend scales were linked to previous assessment scales via a common population linking procedures described in the subject-specific data analysis chapters. Essentially, the 1990 and 1992 data were calibrated together. For each scale, the mean and standard deviation of the 1990 data from this joint calibration were matched to the mean and standard deviation of the 1990 data as previously reported. This then linked the 1992 data to the previously established scale. The writing long-term trend scale was a new IRT-based scale established across four assessments. New scales were also established for the reading and writing main assessments. The metrics for the newly established scales were set to have a mean of 250 and a standard deviation of 50.

The transformations were of the form

$$\theta_{proficiency} = A \bullet \theta_{calibrated} + B$$

where

$\theta_{proficiency}$ = scale level in terms of the system of units of the final proficiency scale used for reporting;

$\theta_{calibrated}$ = scale level in terms of the system of units of the provisional NAEP-BILOG/PARSCALE scale;

A = SD_{new} / $SD_{calibrated}$;

B = M_{new} / $AM_{calibrated}$;

SD_{new} = the estimated or selected standard deviation of the proficiency distribution to be matched;

$SD_{calibrated}$ = the estimated standard deviation of the sample proficiency distribution on the provisional NAEP-BILOG/PARSCALE scale;

M_{new} = the estimated or selected mean of the proficiency distribution to be matched; and

$M_{calibrated}$ = the estimated mean of the sample proficiency distribution on the provisional NAEP-BILOG/PARSCALE scale.

The final transformation coefficients for transforming each provisional scale to the final reporting scale are given in subsequent chapters.

Tables of Proficiency Means and Other Reported Statistics

Proficiencies and trends in proficiencies were reported by age for a variety of reporting categories. Additionally, for trends, the percentages of the students within each of the reporting groups who were at or above the previously defined scale anchor points were reported to provide information about trends in the distribution of achievement within each subject area. The variances of all estimates based on proficiency values included the component due to the error due to the latency of proficiency values of individual students as well as the component due to sampling variability.

9.2.4 Drawing Inferences from the Results

Before discussing an observed difference in means, it is necessary to test whether it is large enough to be statistically significant, or could reasonably have arisen from sampling error alone. If the distributions are normal and the observed difference divided by a weighted function of the two group standard errors exceeds the two-sided critical value 1.96, then a difference that large would arise by chance in only $\alpha = .05$ of cases with these standard errors, in which the true group means were equal.

However, if we wish to make a number of comparisons in the same analysis, say White students versus Black, Hispanic, Asian/Pacific Island, and American Indian students, the probability of finding "significance" by chance for at least one comparison increases with the number (family size) of comparisons. By the Bonferroni inequality, for a family size of 4, the probability of a false positive (Type I error) using $\alpha = .05$ is less than or equal to $4 \times .05 = .20$.

One general method for controlling error rates in multiple comparisons is to apply this inequality and divide α by the family size. Now $\alpha' = .05/4 = .0125$, and using α', the combined probability of one or more errors in the 4 comparisons remains controlled at less than or equal to .05. Note that dividing the probability by n is not the same as multiplying the critical value or the confidence band by n. Indeed, in moving from a family size of 1 to 4, we increase the critical value only from 1.960 to 2.498, a 27.4 percent increase. Doubling the family size again, to 8, increases the critical value to 2.735, an additional 9.5 percent increase. To double the initial critical value to 3.92, the family size would have to be increased to 564.

The power of the tests thus depends on the number of comparisons planned. There may be cases where, before the data is seen, it is determined that only certain comparisons will be conducted. As an example, with the five groups above, interest might lie only in comparing the first group with each of the others (family size 4), rather than comparing all possible pairs of groups (family size 10). This means that some possibly significant differences will not be found or discussed, but the planned comparisons will have greater power to identify real differences when they occur.

In situations with large family sizes and consistent, but small (relative to standard error) differences, the Bonferroni test will declare none of the comparisons significant. In such circumstances, a sign test may be preferable.

For trends extending over several administrations, power is gained by testing least-squares fitted linear and quadratic trends, rather than individual pairs of data points. For example, if the linear regression coefficient is significantly greater than zero, and the quadratic coefficient is not different from zero, the trend over time is positive, even though the Bonferroni test might declare no individual pair of points significantly different.

The linear and quadratic components of the trend in average proficiency for a given subject area and age group were estimated by applying two sets of orthogonal contrasts to the set of average proficiencies by year. The linear component of the trend was estimated by the sum $b_1 = \Sigma c_j X_j$, where the X_j are the proficiency means by year and the c_j are defined so that b_1 corresponds to the slope of an unweighted regression of the proficiency means on the assessment year. The quadratic component was estimated by the sum $b_2 = \Sigma d_j X_j$, where the d_j are orthogonal to the c_j and are defined such that b_2 is the quadratic term in the unweighted regression of the proficiency means on the assessment year and the square of the assessment year. The statistical significance of b_1 and b_2 was evaluated by comparing each estimate to its standard error. The standard error of b_1 was computed as the square root of the sum $\Sigma c_j^2 SE_j^2$, where SE_j is the standard error of X_j. The standard error of b_2 was analogously defined. Tests for linear and quadratic trends allow conclusions about the patterns in the means for several points in time or for several related subgroups of students.

9.2.5 Minimum Sample Sizes for Reporting Subgroup Results

In NAEP reports and data summaries, estimates of quantities such as composite and content area proficiency means, percentages of students at or above the achievement levels, and percentages of students indicating particular levels of background variables (as measured in the student, teacher, and school questionnaires) are reported for the total population as well as for key subgroups determined by the background variables. In some cases, sample sizes were not large enough to permit accurate estimation of proficiency and/or background variable results for one or more of the categories of these variables.

For results to be reported for any subgroup, a minimum sample size of 62 was required. This number was arrived at by determining the sample size required to detect an effect size of 0.5 with a probability of .8 or greater[2] The effect size of 0.5 pertains to the "true" difference in mean proficiency between the subgroup in question and the total population, divided by the standard deviation of proficiency in the total population.

9.2.6 Estimates of Standard Errors with Large Mean Squared Errors

Standard errors of mean proficiencies, proportions, and percentiles play an important role in interpreting subgroup results and comparing the performances of two or more subgroups. The jackknife standard errors reported by NAEP are statistics whose quality depends on certain features of the sample from which the estimate is obtained. In certain cases, typically when the number of students upon which the standard error is based is small or when this group of students all come from a small number of participating schools, the mean squared error associated with the estimated standard errors may be quite large. In the summary reports, estimated standard errors subject to large mean squared errors are followed by the symbol "!".

The magnitude of the mean squared error associated with an estimated standard error for the mean or proportion of a group depends on the coefficient of variation (*CV*) of the estimated size of the population group, denoted as *N*. This coefficient of variation is estimated by:

$$CV(\hat{N}) = \frac{SE(\hat{N})}{\hat{N}}$$

where $\hat{N}$ is a point estimate of N and $SE(\hat{N})$ is the jackknife standard error of $\hat{N}$.

Experience with previous NAEP assessments suggests that when this coefficient exceeds 0.2, the mean squared error of the estimated standard errors of means and proportions based on samples of this size may be quite large. Therefore, the standard errors of means and proportions for all subgroups for which the coefficient of variation of the population size exceeds 0.2 are followed by "!" in the tables of all summary reports. These standard errors, and any confidence intervals or significance tests involving these standard errors, should be interpreted with caution. (Further discussion of this issue can be found in Johnson & Rust, 1992.)

---

[2] A design effect of 2 was assumed for this purpose, implying a sample design-based variance twice that of simple random sampling. This is consistent with previous NAEP experience (Johnson & Rust, 1992).

9.2.7 Teacher Questionnaires

Teachers of students who were in certain of the main assessment samples (varying for the different subject areas) were asked to complete a two-part questionnaire. The first part of the questionnaire pertained to the teacher's background and training. The second part pertained to the procedures used by the teacher for *each class* containing an assessed student. (See Chapter 2 for a description of the teacher questionnaires.)

To analyze the data from the teacher questionnaires with respect to the students' data, each teacher's questionnaire had to be matched to all of the sampled students who were taught by that teacher. In the subsequent chapters two separate match rates are given. The first is the percentage of students that could be matched to both the first and second parts of the teacher questionnaire. For these students, information is available not only about the background and training of their teachers, but also about the methods used in the particular class they attended. The second match rate is the percentage of students that could be matched to the first part of the teacher questionnaire. This match rate is larger because more students could be matched with information about a teacher than with information about the particular class they attended. Note that these match rates only reflect the student-level missing data. They do not reflect the additional missing data due to item-level nonresponse on the part of teachers. Variables derived from the teacher questionnaires were used as reporting variables at the student level and as variables that contributed to conditioning for the appropriate samples.

9.2.8 Dimensionality Analysis

Over the years a number of studies have been conducted in order to seek answers to the question of how many dimensions underlie the various NAEP assessment instruments, and whether there is a sufficiently strong first dimension to support inferences about a composite scale in those subject matter areas in which the cognitive assessment instruments comprise more than one scale (mathematics and reading). Findings from those studies (briefly reviewed below) can, for the most part, be applied to current NAEP instruments because there is a heavy emphasis on keeping the frameworks consistent across years. From time to time, however, changes do take place and in cases where this occurs it is necessary to conduct studies to determine the effects of such changes on the dimensionality of the item pools. For the 1992 assessment such a study was conducted (Carlson, 1993) to determine whether the increasing emphasis on extended constructed-response items that are scored polytomously has any effect on the dimensionality.

Previous Dimensionality Analyses of NAEP Data

NAEP reading assessment data collected during the 1983-84 academic year was studied for dimensionality by Zwick (1986, 1987) who also examined simulated data designed to mirror the NAEP reading item-response data but having known dimensionality. Principal components analysis was applied to both phi and tetrachoric correlation matrices and full information item factor analysis (Bock & Aitkin, 1981; Bock, Gibbons, & Muraki, 1988) implemented in the TESTFACT computer program (Wilson, Wood, & Gibbons, 1991) was applied to portions of the dataset, as was Rosenbaum's (1984, 1985) dimensionality testing procedures. Analysis of the simulated datasets allowed her to determine whether the BIB spiraling design artificially increases dimensionality. Zwick found substantial agreement among the various statistical procedures, and that the results using BIB spiraling were similar to results for complete datasets. Overall she concluded that "it is not unreasonable to treat the data as unidimensional (1987, p. 306)."

The topic of Rock's (1991) investigation was "whether the presently reported subscale scores do span a multidimensional space defined by the content area subscales at each of the three grade levels in mathematics and science (p. 1)." He formed two parcels of items that are homogeneous with respect to content for each subtest of the NAEP mathematics and science tests from the 1990 assessment, and studied their dimensionality using confirmatory factor analysis. The resulting factor intercorrelations averaged across booklets ranged from .86 to .95 in mathematics, and from .94 to .96 in science. Rock's conclusion was that there was little evidence for discriminant validity except for the geometry scale at the eighth-grade level, and that "we are doing little damage in using a composite score in mathematics and science (p. 2)."

A second-order factor model was used by Muthén (1991) in a further analysis of Rock's mathematics data, to examine subgroup differences in dimensionality. Evidence of content-specific variation within subgroups was found, but the average (across seven booklets) percentages of such variation was very small, ranging from essentially zero to 22, and two-thirds of these percentages were smaller than 10.

Carlson and Jirele (1992) used the same full information item factor analysis procedure used by Zwick, and normal harmonic factor analysis (McDonald, 1962, 1967, 1981) as implemented in the NOHARM program (Fraser, 1988) to examine 1990 NAEP mathematics data at three grade levels. Analyses of simulated one-dimensional data were also conducted, and the fit to these data, as measured by the root mean square residual and the Akaike Information Coefficient (Akaike, 1987) was slightly better than that to the real NAEP data. The simulated data were generated using a three-parameter IRT model and a BIB spiraling design like that used in NAEP. Although there was some evidence suggesting more than one dimension in the NAEP data, the strength of the first dimension led the authors to conclude that the data "are sufficiently unidimensional to support the use of a composite scale for describing the NAEP mathematics data, but that there is evidence that two dimensions would better fit the data than one (p. 31)."

Analyses of the 1992 NAEP Cognitive Instruments

As mentioned above, the dimensionality analyses conducted on the 1992 assessment data was directed towards study of the question as to whether dimensionality is affected by inclusion of polytomously scored items on the 1992 NAEP cognitive instruments. It should be noted that several of the procedures used by previous researchers could not be used in this fashion because they do not allow for polytomously scored items.

Carlson (1993) used the LISREL computer program, employing a maximum likelihood parameter estimation technique to study the factor structure of 1992 mathematics and reading cognitive instruments. Results were evaluated, as suggested by McDonald (1981, 1982a, 1982b, 1985), through use of residuals from the fitted model. The square root of the mean squared residual was the statistic used.

It should be noted that in the 1992 NAEP instruments used in this study there is only one polytomously scored item in each block of the BIB spiral. Hence there were only three polytomously scored items in each student's responses. Additionally, at the twelfth grade there was one block of mathematics items that had no polytomously scored item so the students assigned booklets containing that block were only administered two such items. In order to revisit one of the questions studied by Zwick (1986, 1987) complete datasets were simulated as well as datasets using the BIB design.

Mathematics Results. As revealed in the values of the root mean square residual statistics there was no obvious difference in the fit with one, two, or three factors at the twelfth-grade level. At the lower grade levels there was some decrease in the root mean square residual when more factors were fitted, but the increase was so minimal that the data were concluded to be essentially unidimensional. Types of items, one of the primary focuses of the research, do not appear to result in multidimensionality in the context of the types of structures in the NAEP mathematics data. That is, there were only minor differences between one-dimensional solutions and a two-dimensional solutions in which the second dimension is defined by the polytomously scored (or constructed-response) items, and the first by the dichotomously scored (or multiple-choice) items.

Reading Results. Similar analyses were conducted on the 1992 NAEP reading assessment data. In the case of reading the lower correlations in the actual data suggested studying more than one simulated factor structure. Because of the specific blocks assembled into the NAEP reading instruments, the actual data used in this study never included items measuring more than one of the three NAEP reading scales. Each block, however, as pointed out above, consists of a reading passage and several items (9 to 13) about that passage. Hence the multidimensional simulated data were generated as if each passage defined a separate dimension. The correlations among the actual reading scales that were used in generating these multidimensional data, were lower than those among the mathematics scales.

In the actual data, fitting more than one factor had more affect on the size of the root mean square residual statistics and interfactor correlations than was the case in mathematics, at least at the eighth- and twelfth-grade levels. Again, however, there appeared to be little or no effect associated with item type—dichotomously- versus polytomously scored, or multiple-choice versus constructed-response. In the case of simulating a complete data matrix of three dimensions at the twelfth-grade level, the root mean square residual statistic did seem to indicate some lack of fit when one or two dimensions were fitted rather than the three that underlie the generation process. The trend in the actual twelfth-grade data shows less of an effect than in the simulated data, suggesting fewer than three dimensions in the NAEP reading instruments.

Conclusions. The relative sizes of the root mean square residual statistics for the simulated as compared to actual data suggested that lack of fit may be more due to the BIB spiraling design of NAEP than the number of dimensions fitted. Consistent with findings by Zwick (1986, 1987), however, the incomplete design for data collection used in NAEP does not appear to be artificially inflating the dimensionality of the instruments. As might be expected, it was found that the sizes of the root mean square residual statistics for the incomplete simulation condition (a BIB design as in the actual NAEP assessment) were more like those of the real data than those of the case of simulation of a complete data matrix.

9.3 OVERVIEW OF CHAPTERS 10 THROUGH 15

The remaining chapters in Part II of this report are as follows:

Chapter 10: The 1992 National Assessment used a stratified multistage probability sampling design that provided for sampling certain subpopulations at higher rates (see Chapter 3). Because probabilities of selection are not the same for all assessed students, sampling weights must be used in the analysis of NAEP data. Also, in NAEP's complex sample, observations are not independent. As a result, conventional formulas for estimating the sampling variance of statistics are inappropriate. Chapter 10 describes the weighting procedures and methods for estimating sampling variance that are necessitated by NAEP's sample design. Further detail on sampling and weighting procedures is provided in *1992 National Assessment of Educational Progress Sampling and Weighting Procedures, National Assessment, Final Report* (Wallace & Rust, 1993), a report prepared by Westat, Inc., the NAEP subcontractor in charge of sampling.

Chapter 11: A major NAEP innovation introduced by ETS is the reporting of subject-area results in terms of proficiency scales. Scaling methods can be used to summarize results even when students answer different subsets of items. For purposes of summarizing item responses, NAEP developed a scaling technique that has its roots in IRT and in the theories of imputation of missing data. Chapter 11 describes this scaling technique, the underlying theory, and the application of these methods to 1992 NAEP data. The final section of Chapter 11 gives an overview of the NAEP scales that were developed for the 1992 assessment.

Chapter 12: The two main components of the 1992 reading analysis are described in this chapter. First, the reading trend results for the years 1971 through 1990 were extended to include 1992 at ages 9, 13, and 17. The results of the reading trend analysis, which include the percentages of students at or above the reading scale anchor points established in 1984, are reported in *NAEP 1992 Trends in Academic Progress: Achievement of U.S. Students in Science, 1969 to 1992; Mathematics, 1973 to 1992; Reading, 1971 to 1992; Writing, 1984 to 1992* (Mullis, Dossey, Campbell, Gentile, O'Sullivan, & Latham, 1994). In addition, a detailed analysis of the main assessment reading was conducted for grades 4, 8, and 12, including a study of the association between reading proficiency and student background variables. The main assessment analyses are reported in the *NAEP 1992 Reading Report Card for the Nation and the States* (Mullis, Campbell, & Farstrup, 1993). A special study of oral reading, the Integrated Reading Performance Record, was also conducted.

Chapter 13: The trend and main assessment analyses of the mathematics data are detailed in Chapter 13. The results of the trend analysis, which provided links from 1976 through 1992 for ages 9, 13, and 17, are reported in *NAEP 1992 Trends in Academic Progress:*

Achievement of U.S. Students in Science, 1969 to 1992; Mathematics, 1973 to 1992; Reading, 1971 to 1992; Writing, 1984 to 1992 (Mullis, Dossey, Campbell, Gentile, O'Sullivan, & Latham, 1994). A detailed main assessment analysis for grades 4, 8, and 12 in 1992 was also conducted, including an examination of the association of mathematics knowledge with instructional techniques and student background variables. The main assessment results appear in *NAEP 1992 Mathematics Report Card for the Nation and the States* (Mullis, Dossey, Owen, & Phillips, 1993). At grades 4 and 8, background information and data on instructional methods were collected from teachers and the relation of these variables to mathematics proficiency was examined.

Chapter 14: Unlike the reading and mathematics assessments, the science assessment in 1992 consisted of only one component, a long-term trend. The science trend results, which provide a link to 1970, 1973, 1977, 1982, 1986, and 1990, are reported in *NAEP 1992 Trends in Academic Progress: Achievement of U.S. Students in Science, 1969 to 1992; Mathematics, 1973 to 1992; Reading, 1971 to 1992; Writing, 1984 to 1992* (Mullis, Dossey, Campbell, Gentile, O'Sullivan, & Latham, 1994).

Chapter 15: Like the mathematics and reading assessments, the writing analysis consisted of two components, a trend and a main assessment. The writing trend results, which provide a link to 1984, 1988, and 1990 for grades 4, 8, and 12, are reported in *NAEP 1992 Trends in Academic Progress: Achievement of U.S. Students in Science, 1969 to 1992; Mathematics, 1973 to 1992; Reading, 1971 to 1992; Writing, 1984 to 1992* (Mullis, Dossey, Campbell, Gentile, O'Sullivan, & Latham, 1994). The results for the main assessment are reported in the *NAEP 1992 Writing Report Card* (Applebee, Langer, Mullis, Latham, & Gentile, 1994). A special part of the assessment was a sample of writing portfolios gathered from the students in the assessment samples. Results for the portfolio analysis are presented in *Windows into the Classroom: NAEP's 1992 Writing Portfolio Study* (Gentile & Martin-Rehrmann, 1994).

Chapter 10

WEIGHTING PROCEDURES AND ESTIMATION OF SAMPLING VARIANCE[1]

Eugene G. Johnson

Educational Testing Service

Keith F. Rust and Leslie Wallace

Westat, Inc.

As was the case in previous assessments, the 1992 National Assessment used a complex sample design with the goal of securing a sample from which estimates of population and subpopulation characteristics could be obtained with reasonably high precision (as measured by low sampling variability). At the same time, it was necessary that the sample be economically and operationally feasible to obtain. The resulting sample had certain properties that had to be taken into account to ensure valid analyses of the data from the assessment.

The 1992 NAEP sample was obtained through a stratified multistage probability sampling design that included provisions for sampling certain subpopulations at higher rates (see Chapter 3). To account for the differential probabilities of selection, and to allow for adjustments for nonresponse, each student was assigned a sampling weight. Section 10.1 discusses the procedures used to derive these sampling weights.

Another consequence of the NAEP sample design is its effect on the estimation of sampling variability. Because of the effects of cluster selection (students within schools, schools within primary sampling units) and because of the effects of certain adjustments to the sampling weights (nonresponse adjustment and poststratification), observations made on different students cannot be assumed to be independent of one another. As a result, ordinary formulas for the estimation of the variance of sample statistics, based on assumptions of independence, will tend to underestimate the true sampling variability. Section 10.2 discusses the jackknife technique used by NAEP to estimate sampling variability. (The estimation of variability due to imperfect measurement of individual proficiency is discussed in Chapter 11.)

The jackknife technique provides good quality estimates of sampling variability but requires considerable computations. Section 10.3 suggests the use of design effects, combined with conventional variance estimation formulas, as a simple approximation to sampling variability.

---

[1] The statistical programming for this chapter was provided by Bruce Kaplan and Phillip Leung of Educational Testing Service, and Annell Bond, Marianne Whitlock, Fran Cohen, and Lana Ryaboy of Westat, Inc.

Yet another effect of the multistage NAEP sampling scheme is an increase in the variability of variance estimates, as compared with directly drawing independent samples of students (or of schools) without clustering them. Assuming that the distribution of variance estimates is approximately chi-square, the variability can be taken into account by specifying the degrees of freedom of the approximating chi-square variable: the higher the degrees of freedom, the lower the variability of the estimator. In a simple random sample, the degrees of freedom of a variance estimate depend upon the number of subjects and on the distribution of the variable under consideration. In the NAEP design, the degrees of freedom are primarily a function of the number of primary sampling units and the number of strata in the design, rather than the number of subjects, and the distribution of the variable under consideration has less impact. Section 10.4 discusses the degrees of freedom for NAEP jackknife variance estimates.

Since the sample design determines the derivation of the sampling weights and the estimation of sampling variability, it will be helpful to note the key features of the 1992 NAEP sample design. A description of the design appears in Chapter 3, and the various assessment instruments are detailed in Chapter 4.

The 1992 sample was a multistage probability sample consisting of four stages of selection. The first stage of selection, the primary sampling units (PSUs), consisted of counties or groups of counties. The second stage of selection consisted of elementary and secondary schools. The assignment of sessions to sampled schools comprised the third stage of sampling, and the fourth stage involved the selection of students within schools and their assignment to sessions. The probabilities of selection of the first-stage sampling units were proportional to measures of their size, while the probabilities for subsequent stages of selection were such that the overall probabilities of selection of students were approximately uniform, with exceptions for certain subpopulations that were oversampled by design. For the main assessment, schools with relatively high concentrations of Black students and/or Hispanic students were deliberately sampled at twice the normal rate to obtain larger samples of respondents from those subpopulations, in order to increase the precision in the estimation of the characteristics of these subpopulations. Also for the main assessment, nonpublic school students were sampled at three times the normal rate, again so as to increase the precision of estimates for this population subgroup. For all assessment components, students from schools with smaller numbers of eligibles received lower probabilities of selection, as a means of enhancing the cost efficiency of the sample.

The 1992 main assessment includes three student cohorts: students who were *either* in the fourth grade *or* 9 years old; students who were *either* in the eighth grade *or* 13 years old; and students who were *either* in the twelfth grade *or* 17 years old. The main assessment represents two overlapping samples of students, the first of specified grades (of any age) and the second of specified ages (in any grade). Students were age-eligible if they were born in the appropriate calendar year (1982, 1978, or 1974). The main assessment of all age/grade levels was conducted in the spring of 1992.

The full 1992 assessment also includes a number of additional samples that used the age definitions, times of testing, and modes of administration used in previous assessments. These are referred to as long-term trend samples. Because the purpose of these samples was to provide the statistical linkage between the 1992 data and data from previous assessments, they are also referred to as bridge samples.

The full 1992 NAEP assessment thus includes a number of different samples from several populations. Each of these samples has its own set of weights that are to be used to produce estimates about the characteristics of the population addressed by the sample (the target population). The various samples and their target populations are as follows:

The Main Samples of Students. These samples, one for each of the three age/grade levels, were drawn in January through March of 1992, use the calendar-year age definitions, and consist of all students assessed in the main assessment. The target population for each of these samples consists of all students who are in the specified age/grade who were deemed assessable by their school.

For each age/grade, there were four or five distinct assessments, each conducted in distinct session types. Each age/grade conducted print-administered reading, writing, and mathematics assessments, and a special mathematics assessment administered using a paced audio tape. It was necessary to purchase new calculators for the age 9/grade 4 mathematics assessment. Therefore, an additional mathematics assessment was conducted at age 9/grade 4 to serve as a bridge for any difference in performance due to changing calculators.

Because of these administrative arrangements, the weighting procedures were applied separately to four or five distinct subsets of the main sample of assessed students at each age/grade. A set of weights was produced corresponding to each assessment type. Thus in total there were 13 sets of final student weights for the sample of assessed students in the main assessment—one for each age/grade and assessment type.

Long-term Trend Samples for Reading and Writing. These consist of samples comparable to the 1984 main assessment and address the subject areas of reading and writing. The samples were collected by grade and age for age 9/grade 4, age 13/grade 8, and age 17/grade 11, using the age definitions and time of testing from 1984. As in that assessment, print administration was used. Six assessment booklets were administered at each age/grade. The respondents to the combined set of assigned booklets at a given age/grade constitute a representative sample of the population of students who are of the specified grade *or* of the specified age. The respondents to any one of the booklets also constitute a representative sample.

Long-term Trend Samples for Mathematics and Science. These consist of samples comparable to those used for the measurement of trends in 1986. The samples were collected by age only and used the same age definitions and time of testing as in the long-term trend assessment in 1986. As in that assessment, the administration of mathematics and science questions was paced with an audiotape. For ages 9 and 13, three assessment booklets were administered to each age group while two booklets were administered at age 17. The respondents to any one of the booklets assigned to a given age constitute a representative sample of the population of all students of that age.

For purposes of sampling and weighting, the assessment samples are categorized as "tape-administered" or "print-administered" according to whether or not paced audiotapes are to be used in the administration:

1) *Tape-administered samples* are samples that require audiotape pacing in the assessment (the long-term trend assessments in mathematics and science and the special mathematics assessment). For these samples, all students within a particular assessment session receive the same booklet and were paced through at least part of the booklet with an audiotape. These assessment sessions are accordingly referred to as tape-administered sessions.

2) *Print-administered samples* are all main assessment samples and the long-term trend assessments of reading and writing. For these samples, no audiotape pacing was employed and the assessment booklets presented to a particular sample were spiraled through each assessment session (that is, the booklets were systematically interspersed and assigned for testing in that order). These assessment sessions are referred to as print-administered sessions.

10.1 DERIVATION OF THE SAMPLE WEIGHTS

As indicated previously, NAEP uses differential sampling rates, deliberately oversampling certain subpopulations to obtain larger samples of respondents from those subgroups, thereby enhancing the precision of estimates of characteristics of these oversampled subgroups. As a result of the oversampling, these subpopulations, corresponding to students from schools with high concentrations of Black and/or Hispanic students, and from nonpublic schools, are overrepresented in the sample. Lower sampling rates were introduced also for very small schools (those schools with only 1 to 19 eligible students). This reduced level of sampling from small schools was undertaken in an approximately optimum manner as a means of reducing variances per unit of cost. Appropriate estimation of population characteristics must take disproportionate representation into account. This is accomplished by assigning a weight to each respondent, where the weights approximately account for the sample design and reflect the appropriate proportional representation of the various types of individuals in the population.

The weighting procedures for 1992 included computing the student's base weight, the reciprocal of the probability that the student was selected for a particular session. Such weights are those appropriate for deriving estimates from probability samples via the standard Horvitz-Thompson estimator (see Cochran, 1977, section 9A.7). These base weights were adjusted for nonresponse and then subjected to a trimming algorithm to reduce a few excessively large weights. The weights were further adjusted by a poststratification procedure in an effort to reduce the sampling error and certain potential biases of estimates relating to student populations corresponding to several subgroups of the total population. Poststratification was performed by adjusting the weights of the sampled students so that the resulting estimates of the total number of students in a number of specified subgroups of the population corresponded to population totals based on information from the Current Population Survey and Census Bureau estimates of the population (see Wallace & Rust, 1994). The subpopulations were defined in terms of race, ethnicity, geographic region, age, grade, and modal grade status.

The following sections provide an overview of the procedures used to derive the sampling weights. Further details in the derivation of these weights can be found in Wallace and Rust, 1994.

10.1.1 Student Base Weight

The base weight assigned to a student is the reciprocal of the probability that the student was invited to a particular type of assessment session. That probability is the product of four factors:

1) the probability that the PSU was selected;

2) the conditional probability, given the PSU, that the school was selected;

3) the conditional probability, given the sample of schools in a PSU, that the school was allocated the specified type of session; and

4) the conditional probability, given the school, that the student was selected for the specified type of session.

Thus, the base weight for a student may be expressed as the product

$$W_B = PSUWT \bullet SCHWT \bullet SESSWT \bullet STUSCHW$$

where PSUWT, SCHWT, SESSWT, and STUSCHW are, respectively, the reciprocals of the preceding probabilities.

Tables C-1 and C-2 in Appendix C show the distribution of base weights for each of the separate sessions conducted as part of the 1992 assessment. The variations in probabilities of selection, and consequently of weights, were introduced by design, either to increase the effectiveness of the sample in achieving its goals of reporting for various subpopulations, or to achieve increased efficiency per unit of cost.

10.1.2 Adjustment of Base Weights for Nonresponse

The base weight for a selected student was adjusted by three nonresponse factors. The first of these was to adjust for schools that were not contacted for NAEP because they were included in the 1992 National Educational Longitudinal Study (NELS). The second adjustment was needed to account for sessions that were not conducted, either because a contacted school refused or because a cooperating school did not conduct all of its assigned sessions. The third adjustment was needed to adjust for students who were (or should have been) invited to the assessment but did not appear either in the scheduled session or a makeup session. Thus, the nonresponse adjusted weight for a student is of the form

$$W' = W_B \bullet NELSNRF \bullet SESNRF \bullet STUNRF$$

where the nonresponse adjustment factors NELSNRF, SESNRF, and STUNRF are computed as described below. It should be noted that the nonresponse adjustments assume that nonresponse occurs at random within the categories within which adjustments are made (see Little & Rubin, 1987). Some degree of bias could result to the extent that this assumption is false.

10.1.2.1 School Nonresponse Adjustment (SCHNRF)

The NELS nonresponse adjustment was intended to compensate for schools that were not contacted because they were in NELS. These factors were computed separately within a subuniverse (see Chapter 3); that is within one of 18 classes of PSUs within the same region, certainty/noncertainty status, MSA/nonMSA status, and high minority status.

The NELS nonresponse adjustment factor in subuniverse h, $NELSNRF_h$, is given by

$$NELSNRF_h = \frac{\sum_{i \in A_h} PSUWT_{hi} \cdot SCHWT_{hi} \cdot G_{hi}}{\sum_{i \in B_h} PSUWT_{hi} \cdot SCHWT_{hi} \cdot G_{hi}}$$

where

$PSUWT_{hi}$ = the PSU weight for the PSU containing school i from subuniverse h

$SCHWT_{hi}$ = the school weight for school i in subuniverse h;

G_{hi} = the estimated number of age/grade-eligible students in school i in subuniverse h based on QED data (for sessions involving only age-eligible students, the number of age eligibles in each school was used);

set A_h consists of the *original* sample of schools (cooperating, and noncooperating schools, including NELS schools but not substitutes); and

set B_h consists of the original sample of schools except NELS schools.

For a substitute school, $SCHWT_{hi}$ is defined as the school weight of the originally selected school, while the value of G is taken from the substitute school itself. Occasionally subuniverses were combined to form a single nonresponse class. This occurred when the number of non NELS schools from within a subuniverse was small, leading to undue instability in the school nonresponse adjustment factor prior to such collapsing. Subuniverses collapsed together were as similar in nature as possible.

Tables C-3 and C-4 in Appendix C show the distribution of NELS nonresponse adjustment factors for each of the 1992 assessment sessions.

10.1.2.2 Session Nonresponse Adjustment (SESNRF)

Sessions were assigned to schools before cooperation status was final. The session nonresponse adjustment was intended to compensate for session nonresponse due to refusing schools or individual sessions not conducted. These factors were computed separately within classes formed by subuniverse in the long-term trend samples, and by subuniverse and type of control (two levels; public including Bureau of Indian Affairs and Department of Defense, and private including Catholic) in the main samples. Occasionally additional collapsing of classes was necessary, especially for the smaller assessment components. Whenever possible collapsing for session nonresponse was similar to what had been done for school nonresponse.

In PSU h, the session nonresponse adjustment factor $SESNRF_h$ was given by

$$SESNRF_h = \frac{\sum_{i \in B'_h} PSUWT_{hi} \cdot SCHWT_{hi} \cdot NELSNRF_{hi} \cdot SESSWT_{hi} \cdot G_{hi}}{\sum_{i \in C_h} PSUWT_{hi} \cdot SCHWT_{hi} \cdot NELSNRF_{hi} \cdot SESSWT_{hi} \cdot G_{hi}}$$

where

$PSUWT_{hi}$ = the PSU weight for the PSU containing school i from subuniverse h;

$SCHWT_{hi}$ = the school weight for school i in subuniverse h;

$NELSNRF_{hi}$ = the school nonresponse adjustment due to NELS for school i in subuniverse h;

$SESSWT_{hi}$ = the session allocation weight for school i in subuniverse h;

G_{hi} = the estimated number of age/grade-eligible students in school i in subuniverse h in the case of print-administered sessions, and the estimated number of age-eligible students in the case of the tape-administered sessions, to which only age eligibles were invited (the values of G_{hi} were based on QED data);

set B'_h consists of all in-scope originally sampled schools allocated to the specific type of session in subuniverse h (excluding substitutes); and

set C_h consists of all schools allocated to the session type in subuniverse h that ultimately participated (including substitutes).

Tables C-5 and C-6 in Appendix C show the distribution of the session nonresponse adjustment factor for each of the 1992 sample sessions.

10.1.2.3 Student Nonresponse Adjustment (STUNRF)

Student nonresponse adjustment factors were completed separately for each of the assessment session types.

For assessed students in small main sample sessions, for excluded students, and in the long-term trend samples, the student nonresponse adjustment was made separately for classes of students based on subuniverse and modal grade status (at or above modal grade, below modal grade). In the remaining main samples, the adjustment classes were based on subuniverse, modal grade status, and race class (White or Asian, other). In some cases, nonresponse classes were collapsed into one to improve the stability of the adjustment factors. For each class c in session h, the student nonresponse adjustment factor $STUNRF_{hc}$ is computed by

$$STUNRF_{hc} = \frac{\sum_{j(i) \in A'_{hc}} SCHWT_{hi} \cdot NELSNRF_{hi} \cdot SESSWT_{hi} \cdot SESNRF_{hi} \cdot STUSCHW_{hij(i)}}{\sum_{j(i) \in B'_{hc}} SCHWT_{hi} \cdot NELSNRF_{hi} \cdot SESSWT_{hi} \cdot SESNRF_{hi} \cdot STUSCHW_{hij(i)}}$$

where

$SCHWT_{hi}$ = the school weight for school i in session h;

$NELSNRF_{hi}$ = the school nonresponse adjustment factor due to NELS for school i in session h;

$SESSWT_{hi}$ = the session allocation weight for school i in session h;

$SESNRF_{hi}$ = the session nonresponse adjustment factor in school i in session h;

$STUSCHW_{hij}$ = the within-school student weight for student j in school i in session h;

Set A'_{hc} consists of the students in class c who were invited to session h; and

Set B'_{hc} consists of the students in class c who were assessed in session h.

Tables C-7 and C-8 in Appendix C show the distribution of student nonresponse adjustment factors for each of the 1992 assessment sessions.

10.1.2.4 Evaluation of Potential for Bias Resulting from School and Student Nonparticipation

Although school and student nonresponse adjustments are intended to reduce the potential for nonparticipation to bias the assessment results, they cannot completely eliminate this potential bias with certainty. The extent of bias remains unknown, of course, since there are not assessment data for the nonparticipating schools and students.

Some insight can be gained about the potential for residual nonresponse bias, however, by examining the weighted school- and student-level distributions of characteristics known for both participants and nonparticipants, especially for those characteristics known or thought likely to be related to achievement on the assessment. If the distributions for the full sample of schools (or students) without the use of nonresponse adjustments are close to those for the participants with nonresponse adjustments applied, there is reason to be confident that the bias from nonparticipation is small.

There are several school-level characteristics available for both participating and nonparticipating schools. The tables below show the combined impact of nonresponse and of the nonresponse adjustments on the distributions of schools (weighted by the estimated number of eligible students enrolled) and students, by the type of school (public, Catholic, other private) the size of the school, and the urban/rural nature of the county, as measured by the estimated number of eligible students enrolled. Three size classes have been defined for each age class. The data are for the main assessments all session types combined.

Several student-level characteristics are available for both absent and assessed students. The tables that follow show the impact of school nonresponse and nonresponse adjustments, and student nonresponse and nonresponse adjustments on the distributions of eligible students for each age class. The discussion focuses on the writing assessment, since it is the largest. It is assumed that other large assessments would behave similarly to writing. The small assessments are too unstable to tell. The distributions are presented by age category (two levels), grade category (two levels), race category (White, Black, Hispanic, and other), gender, IEP, and LEP. Age is divided into two categories: at or below modal age, and above modal age. Grade is divided into at or above modal grade, and below modal grade.

Table 10-1 shows the weighted marginal distributions of schools for each of the three classification variables for each age class, using the full sample of in-scope schools—those participating, plus those refusals for which no substitute participated. Table 10-2 shows the same distribution based only on participating schools, with school nonresponse adjustments applied to them. For school-level data, the school nonresponse adjustment is actually a composite of the school and session nonresponse adjustment factors derived for use with student-level data.

It can be seen from the tables that, even though the level of school nonparticipation is as high as 18 percent for age class 17 (see Table 3-11), and somewhat lower for the other age classes, the distributions for the three characteristics considered remain similar.

Tables 10-3 and 10-4 present similar data for students. Table 10-3 shows the distributions for assessed and absent students (with base weights adjusted for school nonparticipation) while Table 10-4 shows them for assessed students only, with the student nonresponse adjustments also applied to the weights. The rates of student nonparticipation were 6.7 percent for age class 9, 11.1 percent for age class 13, and 19.3 percent for age class 17 (see Table 3-11). The tables show that for the distributions of type of school attended and sampling descriptor of community (SDOC), the combined effect of student nonparticipation and the subsequent nonresponse adjustments have resulted in very little change in distribution.

Table 10-1

Distribution of Populations of Eligible Students
Based on Full Weighted Sample of Eligible Schools
1992 Main NAEP Samples

Population	Age 9/Grade 4	Age 13/Grade 8	Age 17/Grade 12
Total population	4,982,439	4,589,257	4,225,639
School type			
Catholic	6.9%	6.7%	6.4%
Other private	4.3%	4.6%	3.9%
Public	88.8%	88.7%	89.7%
School size*			
1	12.9%	11.3%	4.4%
2	30.7%	39.4%	47.3%
3	56.4%	49.3%	48.2%
County type (SDOC)**			
Central city 200,000+	35.0%	34.2%	30.1%
Other 200,000+	18.7%	18.0%	20.2%
Other 25,000+	23.7%	22.6%	22.0%
Other	16.9%	19.9%	20.5%
Extreme rural	5.8%	5.3%	7.3%

---

* School size = number of eligible students enrolled:

	1	2	3
Age 9/Grade 4	1-49	50-99	100+
Age 13/Grade 8	1-49	50-299	300+
Age 17/Grade 12	1-99	100-399	400+

** County type is given by the county-level variable "sampling description of community" (SDOC). For a full definition, see Wallace and Rust (1994).

Table 10-2

Distribution of Populations of Eligible Students Based on Weighted Sample of Participating Schools with School Nonresponse Adjustments 1992 Main NAEP Samples

Population	**Age 9/Grade 4**	**Age 13/Grade 8**	**Age 17/Grade 12**
Total population	4,982,439	4,589,257	4,225,639
School type			
Catholic	7.5%	7.2%	6.6%
Other private	3.7%	4.1%	3.4%
Public	88.9%	88.7%	90.0%
School size*			
1	13.5%	11.6%	4.1%
2	30.7%	38.8%	47.7%
3	55.8%	49.6%	48.2%
County type (SDOC)**			
Central city 200,000+	34.5%	35.1%	29.7%
Other 200,000+	19.5%	16.6%	19.4%
Other 25,000+	23.8%	23.4%	23.2%
Other	15.6%	19.3%	20.6%
Extreme rural	6.6%	5.6%	7.2%

---

* School size = number of eligible students enrolled:

	1	2	3
Age 9/Grade 4	1-49	50-99	100+
Age 13/Grade 8	1-49	50-299	300+
Age 17/Grade 12	1-99	100-399	400+

** County type is given by the county-level variable "sampling description of community" (SDOC). For a full definition, see Wallace and Rust (1994).

Table 10-3

Distribution of Populations of Eligible Students Based on Assessed and Absent Students from Participating Schools 1992 Main NAEP Samples

Population	Age 9/Grade 4	Age 13/Grade 8	Age 17/Grade 12
Total population	4,354,268	4,435,233	3,230,717
School type			
Catholic	7.5%	6.3%	7.0%
Other private	3.6%	3.3%	3.7%
Public	88.9%	90.4%	89.3%
County type (SDOC)*			
Central city, 200,000+	30.0%	32.9%	28.5%
Other 200,000+	20.6%	15.4%	19.3%
Other 25,000+	26.0%	23.9%	23.7%
Other	16.6%	20.2%	20.4%
Extreme rural	6.1%	6.8%	7.3%

---

* County type is given by the county-level variable "sampling description of community" (SDOC). For a full definition, see Wallace and Rust (1994).

Table 10-4

Distribution of Populations of Eligible Students Based on Assessed Students from Participating Schools with Student Nonresponse Adjustments Applied 1992 Main NAEP Samples

Population	Age 9/Grade 4	Age 13/Grade 8	Age 17/Grade 12
Total population*	4,354,262	4,435,229	3,230,722
School type			
Catholic	7.7%	6.7%	7.8%
Other private	3.7%	3.5%	3.9%
Public	88.6%	89.5%	88.3%
County type (SDOC)**			
Central city 200,000+	29.9%	32.7%	28.0%
Other 200,000+	20.4%	15.6%	20.0%
Other 25,000+	26.3%	24.1%	23.3%
Other	16.7%	19.9%	20.2%
Extreme rural	6.0%	7.0%	7.6%

---

* Totals do not match those in Table 10-3 due to rounding.

** County type is given by the county-level variable "sampling description of community" (SDOC). For a full definition, see Wallace and Rust (1994).

Tables 10-5, 10-6, and 10-7 show that both school and student-level nonresponse and nonresponse adjustments have little effect on the distributions of eligible students. All of the distributions in the tables are similar. Table 10-5 shows the weighted distributions of eligible students in participating schools. The distributions are calculated using the base weights of invited students unadjusted for school-level nonresponse. Table 10-6 shows the same distributions using weights adjusted for school nonresponse. Table 10-7 shows the weighted distributions of eligible students using the student nonresponse adjusted weights of assessed students.

When comparing Tables 10-6 and 10-7, we expect the distributions by age category, grade category, and race/ethnicity to be similar because these variables were used to determine student nonresponse adjustment classes. However, the distributions by gender, IEP, and LEP are also similar. To the extent that nonrespondents would perform like respondents with the same characteristics (defined by the classification variables in the tables), the bias in the assessment data is small.

Further information about potential nonresponse bias can be gained by studying the absent students. NAEP proficiency estimates are biased to the extent that assessed and absent students within the same weighting class differ in their distribution of proficiency. It seems likely that the assumption that absent students are similar in proficiency to assessed students is reasonable for some absent students—namely, those whose absence can be characterized as random. Conversely, it seems likely that students with longer and more consistent patterns of absenteeism—such as truants, dropouts, near dropouts, and the chronically ill—are unlikely to be as proficient as their assessed counterparts.

In the 1992 assessments, schools were asked to classify each absent student into one of nine categories. The results of this classification for the writing assessment are shown in Table 10-8. The discussion focuses on the writing assessment because it is the largest. It is assumed that the other large assessments would behave similarly to writing. The smaller assessments are too unstable for analysis by classification of absent students.

Table 10-8 shows that, as anticipated, the majority of absence from the assessment was the result of an absence from school of a temporary and unscheduled nature. Past NAEP experience has shown that absence among 17 year olds occurs at over three times the rate of absence among 9 year olds. This is also evident in the table. Table 10-8 also shows little difference across ages in the proportion of absence classified as temporary—84.2 percent for age class 9, 81.2 percent for age class 13 and 80.0 percent for age class 17. This suggests strongly that a substantial proportion of the temporary absences among age class 17 students is not a result of illness, because such absences are occurring at almost three times the rate that they do among age class 9 students. Whereas it might be reasonable to regard temporary absence due to illness as independent of proficiency, for other temporary absences, this appears less tenable. The data in the table give support to the contention that, at age class 9, student absences are unlikely to introduce any significant bias into NAEP estimates. The absentee rate is low; most absences are temporary, and one third of the remaining absences are a result of parental refusal.

At age class 17, however, a somewhat different picture emerges. A significant component of absenteeism is contributed by students who are in school on the day of the

Table 10-5

Distribution of Populations of Eligible Students in Participating Schools
Based on Invited Students from Participating Schools
Using Base Weights Unadjusted for School Nonresponse
1992 Main NAEP Writing Sample

Population	Age 9/Grade 4	Age 13/Grade 8	Age 17/Grade 12
Total population	3,769,885	3,849,554	2,734,548
Age category			
At modal age or younger	70.8%	70.7%	77.1%
Older than modal age	29.2%	29.3%	22.9%
Grade category			
At or above modal grade	74.6%	72.8%	72.0%
Below modal grade	25.4%	27.2%	28.0%
Age category, grade category			
At or younger, at or above	45.3%	43.5%	49.1%
At or young, below	25.4%	27.2%	28.0%
Older, at or above	29.2%	29.3%	22.9%
Race/ethnicity category			
White	69.8%	71.5%	76.8%
Black	13.2%	14.2%	13.4%
Hispanic	14.0%	12.3%	8.7%
Other	3.0%	2.1%	1.1%
Gender			
Male	50.7%	52.3%	51.7%
Female	49.3%	47.7%	48.2%
IEP			
Yes	3.2%	3.9%	1.9%
No	96.8%	96.1%	98.1%
LEP			
Yes	1.2%	1.0%	1.0%
No	98.8%	99.0%	99.0%
IEP, LEP			
IEP yes, LEP yes	0.1%	0.1%	0.1%
IEP yes, LEP no	3.1%	3.8%	1.9%
IEP no, LEP yes	1.2%	0.9%	0.9%
IEP no, LEP no	95.6%	95.2%	97.2%

Table 10-6

Distribution of Populations of Eligible Students
Based on Invited Students in Participating Schools
Using Weights adjusted for School Nonresponse But Not Student Nonresponse
1992 Main NAEP Writing Sample

Population	Age 9/Grade 4	Age 13/Grade 8	Age 17/Grade 12
Total population	4,354,268	4,435,233	3,230,717
Age category			
At modal age or younger	70.8%	70.7%	77.0%
Older than modal age	29.2%	29.3%	23.0%
Grade category			
At or above modal grade	74.6%	72.8%	72.1%
Below modal grade	25.4%	27.2%	27.9%
Age category, grade category			
At or younger, at or above	45.4%	43.5%	49.1%
At or young, below	25.4%	27.2%	27.9%
Older, at or above	29.2%	29.3%	23.0%
Race/ethnicity category			
White	70.2%	71.4%	76.9%
Black	12.6%	14.0%	13.0%
Hispanic	14.1%	12.4%	9.0%
Other	3.1%	2.1%	1.1%
Gender			
Male	50.8%	52.4%	51.6%
Female	49.2%	47.6%	48.2%
IEP			
Yes	3.8%	3.9%	1.9%
No	96.2%	96.1%	98.1%
LEP			
Yes	1.0%	1.0%	1.0%
No	99.0%	99.0%	99.0%
IEP, LEP			
IEP yes, LEP yes	0.1%	0.1%	0.1%
IEP yes, LEP no	3.8%	3.8%	1.8%
IEP no, LEP yes	0.9%	0.9%	1.0%
IEP no, LEP no	95.3%	95.2%	97.1%

Table 10-7

Distribution of Populations of Eligible Students
Based on Assessed Students from Participating Schools
with Student Nonresponse Adjustments Applied
1992 Main NAEP Writing Sample

Population	Age 9/Grade 4	Age 13/Grade 8	Age 17/Grade 12
Total population*	4,354,262	4,435,229	3,230,722
Age category			
At modal age or younger	70.7%	70.6%	76.9%
Older than modal age	29.3%	29.4%	23.1%
Grade category			
At or above modal grade	74.9%	72.9%	72.2%
Below modal grade	25.1%	27.1%	27.8%
Age category, grade category			
At or younger, at or above	45.6%	43.6%	49.1%
At or young, below	25.1%	27.1%	27.8%
Older, at or above	29.3%	29.4%	23.1%
Race/ethnicity category			
White	70.2%	71.4%	76.9%
Black	12.3%	13.6%	12.7%
Hispanic	14.2%	12.8%	9.2%
Other	3.2%	2.2%	1.2%
Gender			
Male	50.9%	52.3%	51.5%
Female	49.1%	47.7%	48.5%
IEP			
Yes	3.1%	3.6%	1.8%
No	96.9%	96.4%	98.2%
LEP			
Yes	1.2%	0.9%	1.1%
No	98.8%	99.1%	98.9%
IEP, LEP			
IEP yes, LEP yes	0.0%	0.1%	0.1%
IEP yes, LEP no	3.0%	3.5%	1.7%
IEP no, LEP yes	1.2%	0.9%	1.1%
IEP no, LEP no	95.8%	95.5%	97.2%

---

* Figures do not match those in Table 10-6 due to rounding.

Table 10-8

Weighted Distribution of Absent Students by Nature of Absenteeism
for All Age Classes, 1992 Writing Assessment

Nature of Absenteeism	Age 9/Grade 4	Age 13/Grade 8	Age 17/Grade 12
Temporary absence*	84.2%	81.2%	80.0%
Long-term absence**	0.6%	0.9%	1.0%
Chronic truant	0.0%	1.0%	3.0%
Suspended or expelled	0.2%	2.5%	0.5%
Parent refusal	4.8%	2.1%	1.0%
Student refusal	0.1%	0.2%	2.4%
In school, did not attend session	1.3%	2.8%	9.2%
In school, not invited***	3.5%	0.7%	1.0%
Other	3.9%	7.9%	1.5%
Missing	1.2%	0.6%	0.3%
Total absentee sample	687	1,906	3,713
Total sample size	10,239	16,848	19,382
Overall absentee rate	6.7%	11.1%	19.3%

---

* Absent less than two weeks due to illness, disability, or excused absence.
** Absent more than two weeks due to illness or disability.
*** In school, but not invited to assessment session due to disruptive behavior.

assessment but failed to attend the session (9.2%). Chronic truants, those suspended, and those in school but not invited, constitute the obvious candidates for potential bias. As they comprise only 4.5 percent of absent students, their potential for introducing significant bias under the current procedures is minor.

10.1.3 Trimming of Weights

In a number of cases, students were assigned relatively large weights. One cause of large weights was underestimation of the number of eligible students in some schools leading to inappropriately low probabilities of selection for those schools. A second major cause is the presence of large schools (high schools in particular) in PSUs with small selection probabilities. In such cases, the maximum permissible within-school sampling rate (determined by the maximum sample size allowed per school—see Chapter 3) could well be smaller than the desired overall within-PSU sampling rate for students. Large weights arose also because very small schools were, by design, sampled with low probabilities. Other large weights arose as the result of high levels of nonresponse coupled with low to moderate probabilities of selection, and the compounding of nonresponse adjustments at various levels.

Students with notably large weights have an unusually large impact on estimates such as weighted means. Since, under some simplifying assumptions, the variability in weights contributes to the variance of an overall estimate by an approximate factor $1 + V^2$, where V^2 is the relative variance of the weights, an occasional unusually large weight is likely to produce large sampling variances of the statistics of interest, especially when the large weights are associated with students with atypical performance characteristics.

To reduce this problem, a procedure of trimming a few of the more extreme weights to values somewhat closer to the mean weight was applied. This trimming can increase the accuracy of the resulting survey estimates, substantially reducing V^2 and hence the sampling variance while introducing a small bias. The trimming algorithm was identical to that used in the 1984, 1986, 1988, and 1990 assessments and had the effect of trimming the weights of students from any school that contributed more than a specified proportion, ζ, to the estimated variance of the estimated number of students eligible for assessment. The trimming was done separately for each assessment session type. In each case, the value of the proportion ζ was chosen to be 10/K, where K was the number of schools in which a specified assessment was conducted. The number of schools where weights were trimmed was small in each of the samples. Tables C-9 and C-10 in Appendix C show the distribution of trimming factors for each of the 1992 assessment sessions. From the table it is seen that the most extreme trimming factors applied were of the order of 0.4 (except for one that was 0.2) and that trimming affects the weights of only a very small proportion of the assessed students.

Table 10-9 shows the distributions of eligible students based on the trimmed weights of assessed students for the writing samples for each age class. The distributions are similar to those before trimming shown in Tables 10-5, 10-6, and 10-7. To the extent that the characteristics in the table are related to student performance on the writing assessment, the bias in the assessment results introduced by trimming is very small.

Table 10-9

Distribution of Populations of Eligible Students
Based on Trimmed Weights of Assessed Students in Participating Schools
1992 Main NAEP Writing Sample

Population	Age 9/Grade 4	Age 13/Grade 8	Age 17/Grade 12
Total population	4,316,685	4,378,199	3,155,810
Age category			
At modal age or younger	70.6%	70.7%	76.9%
Older than modal age	29.4%	29.3%	23.1%
Grade category			
At or above modal grade	75.1%	72.9%	72.3%
Below modal grade	24.9%	27.1%	27.8%
Age category, grade category			
At or younger, at or above	45.8%	43.7%	49.2%
At or young, below	24.9%	27.1%	27.7%
Older, at or above	29.4%	29.3%	23.1%
Race/ethnicity category			
White	70.3%	71.1%	76.7%
Black	12.4%	13.7%	12.8%
Hispanic	14.3%	13.0%	9.3%
Other	3.0%	2.2%	1.2%
Sex			
Male	51.0%	52.2%	51.4%
Female	49.0%	47.8%	48.5%
IEP			
Yes	3.1%	3.6%	1.8%
No	96.9%	96.4%	98.2%
LEP			
Yes	1.2%	0.9%	1.2%
No	98.8%	99.1%	98.8%
IEP, LEP			
IEP yes, LEP yes	0.0%	0.1%	0.1%
IEP yes, LEP no	3.1%	3.5%	1.7%
IEP no, LEP yes	1.2%	0.9%	1.1%
IEP no, LEP no	95.7%	95.5%	97.1%

10.1.4 Poststratification

As in most sample surveys, the respondent weights are random variables that are subject to sampling variability. Even if there were no nonresponse, the respondent weights would at best provide unbiased estimates of the various subgroup proportions. However, since unbiasedness refers to average performance over a conceptually infinite number of replications of the sampling, it is unlikely that any given estimate, based on the achieved sample, will exactly equal the population value. Furthermore, the respondent weights have been adjusted for nonresponse and a few extreme weights have been reduced in size.

To reduce the mean squared error of estimates using the sampling weights, these weights were further adjusted so that estimated population totals for a number of specified subgroups of the population, based on the sum of weights of students of the specified type, were the same as presumably better estimates based on composites of estimates from the 1989 and 1990 Current Population Survey and 1992 population projections made by the Census Bureau. This adjustment, called poststratification, is intended especially to reduce the mean squared error of estimates relating to student populations that span several subgroups of the population, and thus to reduce the variance of measures of changes over time for such student populations.

10.1.4.1 1992 Poststratification Procedures

The poststratification in 1992 was done for each age/grade and separately for each of the print-administered assessments and each of the groups of tape-administered assessments. Within each age/grade and assessment type group, poststratification adjustment cells were defined in terms of race, ethnicity, and NAEP region as shown in Table 10-10.

Table 10-10
Major Subgroups for Poststratification in 1992

Subgroup	Race	Ethnicity	Region*
1	White	Not Hispanic	Northeast
2	White	Not Hispanic	Southeast
3	White	Not Hispanic	Central
4	White	Not Hispanic	West
5	Any	Hispanic	Any
6	Black	Not Hispanic	Any
7	Other	Not Hispanic	Any

---

* Regions are the same as for stratification and reporting (see Chapter 3), except that all of Virginia is included in the Southeast region for poststratification purposes.

The result is seven poststratification cells for each group of tape-administered sessions. For the assessments involving both age- and grade-eligible students, each of the seven subgroups was further divided into two or three eligibility classes. For age classes 9 and 13 and for the age 17/grade 11 long-term trend sample, three eligibility classes were used:

a) students eligible by grade and of modal age;
b) students eligible by age only;
c) students eligible by grade but not of modal age.

For the age 17/grade 12 main assessment sample, each of the seven subgroups was divided into two subclasses:

a) students eligible by grade (of any age);
b) students eligible by age only.

This variation in the procedure from that used for the other age classes and for the age 17/grade 11 long-term trend sample was adopted because the independent estimates of the numbers of students in the population did not provide consistent data on the numbers of twelfth grade students eligible only by grade. This procedure for age 17/grade 12 is identical to that employed for the 1988 (and 1990) assessment. (See Rust, Bethel, Burke, & Hansen, 1990, 1992, for further details.)

Thus, there were 7, 14, or 21 cells for poststratification. The poststratified weight for each student within a particular cell was the student's base weight, with adjustments for nonresponse and trimming, times a poststratification factor. For each cell, the poststratification factor is a ratio whose denominator is the sum of the weights (after adjustments for nonresponse and trimming) of assessed and excluded students, and whose numerator is an adjusted estimate of the total number of students in the population who are members of the cell. This estimated total was a composite based on the October 1989 and 1990 Current Population Surveys and 1992 population projections. Details of the procedures used to obtain these composite independent estimates are provided in Wallace and Rust (1994). Tables C-11 and C-12 in Appendix C show the distribution of poststratification factors for each of the 1992 assessments.

10.1.4.2 Differences from Earlier Procedures

The poststratification procedure utilized for 1992 was exactly like that used in 1990, and was similar to that used in 1988. This differed somewhat from the procedures used in 1986 and 1984, and the nature and impact of these differences are discussed in Johnson and Zwick (1990, section 10.1.4). There were two differences from the 1988 procedures. The first was in the definitions of the samples that were to be poststratified to the appropriate population totals for the assessments involving only age eligible students. In 1988 each individual tape-administered session was separately poststratified (these are the session types where only age-eligible students are assessed), whereas in 1990 and 1992 these sessions were formed into groups of sessions which essentially constituted a single assessment (two groups at each age), and these groups were poststratified (and indeed weighted in general) as a whole. Conversely, as in 1988, the print-administered sessions were poststratified individually, although in most cases a number of different booklets were administered in a single session.

This change from 1988 was designed to provide a consistent approach whereby all the assessment components that were administered distinctly and analyzed together were poststratified together to the total population. This avoided the potential problems of increased variance that arise when small individual session types are poststratified, with resulting high sampling variability in the poststratification factors derived.

The second difference affected those assessment components to which both age- and grade-eligible students were invited, for age 9/grade 4 and age 13/grade 8. In both the 1988 and the 1992 assessments, students in the modal grade were classified into two groups on the basis of age. In 1988 this split was determined by whether or not the student was age eligible (in addition to being grade eligible). In 1992 the split was determined by whether or not the student was of the modal age (in years) for the grade as of October 1, 1991. This change was made because the independent estimates of these two components of the population for each grade could be made more reliably under the new procedure. This is because the 1992 classification was consistent with the data available from the Current Population Survey estimates, whereas in 1988 a modification was needed, which was based on certain assumptions about the joint distribution of the population by grade and age. These assumptions were in general supported by the available data, but were of necessity simplifying in nature, and therefore a potential source for a small amount of error. Full details of the 1988 procedure for obtaining the independent estimates are given in Appendix C of Rust et al. (1990).

For the age 17/grade 11 long-term trend print-administered sessions and the age 17/grade 12 main assessment samples, no changes were made from the 1988 procedures. At age 17/grade 11, the definition of age eligibility coincided with the modal age as of October 1, 1991, so that there was no change necessary. For age 17/grade 12, there was no split of the grade 12 students by age for purposes of poststratification, in 1988, 1990 or 1992, as discussed above and in Appendix D of Rust et al. (1990).

10.1.5 The Final Student Weight

The final weight assigned to a student is the student full-sample weight. This weight is the student's base weight after the application of the various adjustments described above. The student full-sample weight was used to derive all estimates of population and subpopulation characteristics that have been presented in the various NAEP reports, including simple estimates such as the proportion of students of a specified type who would respond in a certain way to an item and more complex estimates such as mean proficiency levels.

The effects of all of the adjustments to the base weights are summarized in Tables C-13 and C-14 in Appendix C, which show the distribution of the single factor given as the product of NELSNRF, SESNRF, STUNRF, the trimming factor, and the poststratification factor, for each of the assessment components. The distributions of the final student weights are given in Tables C-15 and C-16 in Appendix C.

As indicated earlier, under some simplifying assumptions the factor $1 + V^2$ indicates the approximate relative increase in variance of estimates resulting from the variability in the weights. The factor $1 + V^2$ for each sample is readily derivable from Tables C-15 and C-16 by adding 1 to the square of the ratio of the standard deviation to the mean weight. These factors, resulting from the combined effect of the variations in weights introduced by design and from other causes, are discussed above.

10.1.6 Other Weights

In addition to the weights for the assessed students, weights were also derived for excluded students. In addition, special weighting adjustments were developed for certain subsets of the fourth grade and eighth-grade students assessed in the main samples. The weights for these subsets, with these adjustments applied, were used in equating the results of the national assessment and the Trial State Assessments for subjects they had in common. Finally, a set of weights appropriate for analyzing school-level data files was developed.

Weights for excluded students. Excluded students are students who were designated by the schools as unable to complete the assessment because they were non-English speaking, mildly mentally retarded (educable), or functionally disabled. Since the same grade and age eligibility definitions apply, no distinction is made between students excluded from the various sessions within an assessment. However, the excluded students from the long-term trend and main assessments were treated as two separate samples of excluded students for each age class. This was in part because the guidelines to school personnel for excluding students differed between the main and long-term trend samples, so that the excluded student populations may have differed between these assessment types. The distribution of the final weights for excluded students and the components of the weights are included in Tables C-1 to C-16 in Appendix C.

For the long-term trend samples, students could be potentially excluded from a tape-administered session for which only age-eligible students were selected, or a print-administered session, for which both grade- and age- eligible students were selected. The samples of excluded students from the long-term trend assessments were weighted to reflect the full grade- and/or age-eligible population. This was achieved by weighting each grade-only eligible (i.e., not eligible by age) student who was excluded from a print-administered session to account for his/her probability of assignment to a print-administered session. No such corresponding session assignment adjustment was needed for the age-eligible excluded students, since they were eligible to have been selected for any of the long-term trend assessment session types.

As in the case of the weights for the assessed students, the excluded student weights were constructed from components reflecting the probability of selection, correction for nonresponse, weight trimming, and poststratification. Further details on the derivation of the excluded student weights can be found in Wallace and Rust (1993).

Weights for equating national and state-by-state assessments. The eighth-grade mathematics assessment and the fourth-grade mathematics and reading assessments conducted in February 1992 in each of 41 states, the District of Columbia, and two territories consisted of identical assessment material to that administered in the corresponding main sample sessions. Technical details of the Trial State Assessment Program are given in Johnson, Mazzeo, and Kline (1993). The national and state-by-state assessments were equated so that state and national results could be reported on a common scale. The equating was achieved by using from each assessment that part of the sample representing a common population. For the age 13/grade 8 national sample, this consisted of those eighth-grade public-school students from a participating state (including the District of Columbia) who were assessed in the main sample mathematics assessment.

Although each sample of students received appropriate weights from the weighting procedure used for the national assessment, in an effort to increase the precision of the equating process, an additional weighting adjustment was developed and applied to this subsample, solely for use in equating. The adjustment involved adjusting the distributions of the weights for three categorical variables to agree closely with those obtained from the weighted aggregate sample from the state assessments from the participating states. The first two variables were region (Northeast, Southeast, Central, and West) and race/ethnicity (White nonHispanic, Black nonHispanic, Hispanic, and "other"). The third variable was type of mathematics course taken (algebra, pre-algebra, eighth-grade mathematics, and "other") for eighth-grade mathematics, I am good at math (agree, undecided, disagree) for fourth-grade mathematics, and kind of reader (very good, good, average, poor) for fourth-grade reading. The equating of the weight distribution was achieved using a procedure known as Iterative Proportional Fitting (IPF), described in Little and Rubin, 1987. Raking adjustments were applied to the national sample weights to force their distribution to agree with that from the aggregated state samples, for each of these three variables in turn. This process was then repeated, and the final set of adjusted weights was compared with the state sample weights on all three distributions, and found to be in very close agreement. The resulting adjustments to the national weights ranged in magnitude from a factor of 0.61 to 1.92.

School weights. The sampling procedures used to obtain national probability samples of assessed students also gave rise indirectly to several national probability samples of schools (from which the students were subsequently sampled). So that the school samples can be utilized for making national estimates about schools, appropriate nonresponse adjusted survey weights have been developed.

The weight for each school is partly composed of a base weight, giving the inverse of the selection probability of the school. This weight, W_{BS}, is given by

$$W_{BS} = PSUWT \bullet SCHWT$$

School nonresponse adjustments were then applied to these base weights. These are very similar to the school nonresponse adjustment factors used for student weights, NELSNRF, and were created using the same set of nonresponse adjustment classes. The values of the adjustment factors are not the same, however. A school that was assigned a proper subset of the possible assessment sessions for a given assessment but did not participate at all was treated as not responding at the session level for the student weighting (since its nonparticipation did not affect those session types that were not assigned to it). Such a school was treated as a nonresponding school in creating the school weights.

A total of six samples of schools were weighted to be nationally representative. At each age/grade level, there were two such samples, one being the sample of schools selected for the long-term trend assessment, and the second being the sample of schools selected for the main assessment. At age 9/grade 4, the population of schools represented in each case consists of all schools having at least one of the grades 2 through 5. The school population at age 13/grade 8 is that of schools having at least one of the grades 6 through 9, while the school population for both age 17/grade 11 and age 17/grade 12 is that of schools having at least one of the grades 9 through 12.

Jackknife replicate weights. In addition to the weights that were used to derive all estimates of population and subpopulation characteristics, other sets of weights, called jackknife replicate weights, were derived to facilitate the estimation of sampling variability by the jackknife variance estimation technique. These weights and the jackknife estimator are discussed in the next section.

10.2 PROCEDURES USED BY NAEP TO ESTIMATE SAMPLING VARIABILITY

A major source of uncertainty in the estimation of the value in the population of a variable of interest exists because information about the variable is obtained on only a sample from the population. To reflect this fact, it is important to attach to any statistic (e.g., a mean) an estimate of the sampling variability to be expected for that statistic. Estimates of sampling variability provide information about how much the value of a given statistic would be likely to change if the statistic had been based on another, equivalent, sample of individuals drawn in exactly the same manner as the achieved sample.

Another important source of variability is that due to imprecision in the measurement of individual proficiencies. For the 1992 assessment, proficiencies in all subject areas were summarized through item response theory (IRT) models, but not in the way that these models are used in standard applications where each person responds to enough items to allow for precise estimation of that person's proficiency. In NAEP, each individual responds to relatively few items so that individual proficiency values are not well determined. Consequently, the variance of any statistic based on proficiency values has a component due to the imprecision in the measurement of the proficiencies of the sampled individuals in addition to a component measuring sampling variability. The estimation of the component of variability due to measurement imprecision and its effect on the total variability of statistics based on proficiency values are discussed in Chapter 11.

The estimation of the sampling variability of any statistic must take into account the sample design. In particular, because of the effects of cluster selection (students within schools, schools within PSUs) and because of effects of nonresponse and poststratification adjustments, observations made on different students cannot be assumed to be independent of each other (and are, in fact, generally positively correlated). Furthermore, to account for the differential probabilities of selection (and the various adjustments), each student has an associated sampling weight, which should be used in the computation of any statistic and which is itself subject to sampling variability. Ignoring the special characteristics of the sample design and treating the data as if the observations were independent and identically distributed, will generally produce underestimates of the true sampling variability.

The proper estimation of the sampling variability of a statistic based on the NAEP data is complicated and requires techniques beyond those commonly available in standard statistical packages. Fortunately, the *jackknife* procedure (see, e.g., Wolter, 1985; Kish & Frankel, 1974; Rust, 1985) provides good quality estimates of the sampling variability of most statistics, at the expense of increased computation, and can be used in concert with standard statistical packages to obtain a proper estimate of sampling variability.

The jackknife procedure used by NAEP has a number of properties that make it particularly suited for the analysis of NAEP data. When properly applied, a jackknife estimate of the variability of a linear estimator (such as a total) will be the same as the standard textbook variance estimate specified for the sample design (if the first-stage units were sampled with replacement and approximately so otherwise). Additionally, if the finite sampling corrections for the first stage units can be ignored, the jackknife produces asymptotically consistent variance estimates for statistics such as ratios, regression estimates or weighted means and for any other nonlinear statistic that can be expressed as a smooth function of estimated totals of one or more variables (Krewski & Rao, 1981).

Through the creation of student replicate weights (defined below), the jackknife procedure allows the measurement of variability attributable to the use of poststratification and other weight adjustment factors that are dependent upon the observed sample data. Once these replicate weights are derived, it is a straightforward matter to obtain the jackknife variance estimate of any statistic.

The jackknife procedure in this application is based upon the development of a set of 56 jackknife replicate weights for each assessed student (or excluded student, or school depending upon the file involved). The 56 replicate weights are developed in such a way that, when utilized as described below, approximately unbiased estimates of the sampling variance of an estimate result, with an adequate number of degrees of freedom to be useful for purposes of making inferences about the parameter of interest. For a discussion of the degrees of freedom for variance estimation, see section 10.4.

The estimated sampling variance of a parameter estimator t is the sum of 56 squared differences:

$$\hat{Var}(t) = \sum_{i=1}^{56} (t_i - t)^2$$

where t_i denote the estimator of the parameter of interest, obtained using the ith set of replicate weights, $SRWT_i$, in place of the original set of full sample estimates WT_i. The methods for deriving these replicate weights, $SRWT_i$, are outlined below and full details are given in Wallace and Rust (1993).

Of the 56 replicate weights formed for each record, 30 act to reflect the amount of sampling variance contributed by the noncertainty strata of PSUs, with the remaining 26 replicate weights reflecting the variance contribution of the certainty PSU samples.

The derivation of the 30 replicate weights reflecting the variance of the noncertainty PSUs involves first defining pairs of PSUs (or appropriate aggregates of them in some strata) in a manner that models the design as one in which two PSUs are drawn with replacement per stratum. This definition of pairs is undertaken in a manner closely reflective of the actual design, in that PSUs are paired that are drawn from strata within the same subuniverse, and with similar stratum characteristics. The same definition of pairs was used for each of the assessment components, since all were drawn from the same sample of noncertainty PSUs. The 63 noncertainty PSUs, drawn from 60 strata, were formed into 30 pairs of PSUs, where the pairs were composed of PSUs from adjacent strata within each subuniverse (thus the strata were relatively similar on the characteristics of proportion minority population, population change between 1970 and 1980, and the proportions of urban and farm populations). For those three strata where two PSUs were included in the sample, in each case both PSUs were treated together as constituting a half of one pair. Whereas the actual sample design was to select one PSU with probability proportional to size from each of 60 strata, and then to select supplementary PSUs as needed, for variance estimation purposes the design is regarded as calling for the selection of two PSUs with probability proportional to size with replacement from each of 30 strata. This procedure likely gives a small positive bias to estimates of sampling error.

The student replicate weight for the i^{th} pair of noncertainty PSUs, for the 30 pairs corresponding to values of i from 27 to 56, is computed as follows:

1) Let W_B be the base weight of a student, as described in section 10.1.1, which accounts for the various components of the selection probability for the student.

2) At random, one PSU (or set of PSUs from the same stratum) in each pair is denoted as PSU number 1, while the other is denoted as PSU number 2. The i^{th} replicate base weight, W_{Bi}, is given by:

$$W_{Bi} = \begin{cases} 0 & \text{if the student belongs to PSU number 1 of pair i} \\ 2 * W_B & \text{if the student belongs to PSU number 2 of pair i} \\ W_B & \text{if the student is from neither PSU in pair i} \end{cases}$$

3) The i^{th} student replicate weight $SRWT_i$ is obtained by applying the various school and student nonresponse adjustments, the weight trimming, and the poststratification to the i^{th} set of replicate base weights, using procedures identical to those used to obtain the final student weights WT from the set of base weights W_B.

In brief, the procedure for deriving the sets of W_{Bi} value from the W_B values reflects the sampling of PSUs, schools, sessions, and students. By repeating the various weight adjustment procedures in each set of replicate base weights, the impact of these procedures on the sampling variance of the estimator t is appropriately reflected in the variance estimator $\hat{Var}(t)$ defined above.

The procedure for obtaining the 26 sets of replicate weights to estimate the sampling variance from the certainty PSUs is analogous, but somewhat more complex. The first stage of sampling in this case is at the school level, and the derivation of replicate weights must reflect appropriately the sampling of schools within certainty PSUs. Since each of six different sample components (three age/grade classes by main assessment or long-term trend) involved different samples of schools, the procedure for forming replicate base weights was individualized to each of these six sample components. In common across these six samples were the 34 certainty PSUs used, and the fact that 26 replicate weights were formed in each case.

For a given sample, the 34 certainty PSUs constituted strata, with a sample of schools drawn systematically within each. Using the schools listed in order of sample selection within each stratum, successive schools were paired or formed into triples. These pairs and triples numbered more than 26, so that each replicate weight was in general formed by perturbing the weights of students from more than a single pair or triple. These aggregates of pairs and triples were in general assigned in proportion to the size of the PSU. Thus generally speaking, the four largest PSUs were assigned two replicates each, the next six largest one replicate each, and the remaining 24 were paired and assigned 12 replicates. When splitting the larger PSUs, the schools were split into two groups of (as close as possible) equal size, based on the ordering at the time of sample selection. The first half of the sample was assigned to one replicate, the second half to another. Within each PSU (or half PSU in the case of the four large split PSUs) schools were alternately numbered 1 or 2 starting randomly. If, however, there were exactly three schools sampled in the PSU the schools were randomly numbered 1, 2, or 3. The method of forming replicate base weights in strata where there were not exactly three schools was the same as for the noncertainty strata (except that members of a pair i could come from more than a single "stratum"). When a stratum contained three schools, students in these schools had their weights perturbed for two sets of replicates, say i1 and i2, as follows:

$$W_{Bi1} = \begin{cases} 0 & \text{if the student is in school number 1 of PSU belonging to set i} \\ 1.5 * W_B & \text{if the student is in school number 2 or 3 of a PSU belonging to set i} \\ W_B & \text{if the student does not belong to set i} \end{cases}$$

$$W_{Bi2} = \begin{cases} 1.5 * W_B & \text{if the student is in school number 1 or 2 belonging to set i} \\ 0 & \text{if the student is in school number 3 belonging to set i} \\ W_B & \text{if the student does not belong to set i} \end{cases}$$

The actual pattern of replicate base weight assignment used for each of the nine samples is given in Wallace and Rust (1993).

The nonresponse, trimming, and poststratification adjustments were applied to each set of replicate base weights to derive the final replicate weights in each case, exactly as in the noncertainty PSUs. In fact these procedures were applied to the full set of weights from all parts of the given sample together, just as for the full sample weights. That is, for example, poststratification factors were derived from the full set of data for each replicate, not separately for certainty and noncertainty PSUs.

This estimation technique was used by NAEP to estimate all sampling errors presented in the various reports. A further discussion of the variance estimation procedure used by NAEP, including a discussion of alternative jackknife estimators that were also considered, appears in Johnson (1989).

We noted above (as discussed in Chapter 11) that a separate estimate of the contribution to variance due to the imprecision in the measure of individual proficiencies is made and added to the jackknife estimate of variance. That variance component could have been appropriately reflected in the jackknife variance estimates simply by separately applying the IRT computations to each jackknife replicate. Because of the heavier IRT computational load, this was not done. Less work was involved by the simple procedure of making separate estimates of this component to be added to the jackknife variance estimates. Also, a separate measure of this component of variance is then available, which would not be so if it were reflected in the jackknife variance estimate.

10.3 APPROXIMATING THE SAMPLING VARIANCE USING DESIGN EFFECTS

In practical terms, the major expenditure of resources in the computation of a jackknife variance estimate occurs in the preparation of estimates for each of the pseudoreplicates. In the 1992 assessment, this implies that the statistic of interest has to be recomputed up to 57 times, once for the overall estimate t, and once for each of the up to 56 pseudoreplicates t_i. Because this is a considerable increase in the amount of computation required, relative to a conventional variance estimate, it is of interest to see how much the jackknife variance estimates differ from their less computationally intensive, simple random sampling based, analogues.

The comparison of the conventional and the jackknife methods of variance estimation will be in terms of a statistic called the *design effect*, which was developed by Kish (1965) and extended by Kish and Frankel (1974). The design effect for a statistic is the ratio of the actual variance of the statistic (taking the sample design into account) over the conventional variance estimate based on a simple random sample with the same number of elements. The design effect is the inflation factor to be applied to the conventional variance estimate in order to adjust error estimates based on simple random sampling assumptions to account approximately for the effect of the sample design. The value of the design effect depends on the type of statistic computed and the variables considered in a particular analysis as well as the combined clustering, stratification, and weighting effects occurring among sampled elements. Generally, the design effects for statistics from complex samples such as NAEP are greater than one, because variances based on simple random sampling assumptions tend to provide underestimates of the variances of statistics calculated from complex samples.

10.3.1 Design Effects for Proportion-correct Statistics

As an example of the distribution of design effects to be expected from NAEP data, we consider the design effect for the statistic P, the estimated proportion of a specified subgroup of the population who would correctly respond to a given assessment item. The proportion-correct statistic is the weighted mean of the responses to the item of the assessed individuals who belong to the subgroup, where an individual's response is either 1=correct or 0=incorrect. The design effect for the proportion-correct statistic P is of the form

$$\mathrm{deff}(P) = [\mathrm{Var}_{JK}(P)]/[P(1 - P)/N].$$

In the above, N is the total number of individuals in the subgroup responding to the item, $\mathrm{Var}_{JK}(P)$ is the jackknife variance of P, and P(1 - P)/N is the conventional variance estimate of P. (Although the estimate P(1 - P)/N has the same form as the simple random sampling estimator of the variance of a proportion correct, the use of sample weights in the estimation of P reflects the appropriate distribution of the population.)

The design effects for the proportion-correct statistics for each item administered in a sample are summarized by the mean, median, bottom quartile, top quartile, and standard deviation across proportion-correct statistics for each item in the sample. These summaries are given for each sample in Appendix D. The tables in Appendix D also contain degrees of freedom estimates explained in section 10.4. The numerous tables in Appendix D have been further summarized in Tables 10-11 through 10-19 by averaging the entries in Appendix D across all main samples to produce one summary table for each age/grade level and likewise for all long-term trend and special mathematics samples. For comparison with the national sample results, Tables 10-20 and 10-21 summarize the design effects for the 1992 Trial State Assessment in mathematics and Table 10-22 summarizes the design effects for the Trial State Assessment in reading. The particular demographic variables shown were selected because (1) they are major variables in NAEP reports and (2) they reflect different types of divisions of the population that might have different levels of sampling variability.

There are some systematic differences in the design effects for different types of samples and different subpopulations. The long-term trend samples that were given the tape-administered sessions have larger design effects than the main assessment samples and the long-term trend samples that were print-administered. Because the same items must be administered to each student in a tape-administered session, the number of students per school administered an item is higher for samples with tape-administered sessions than for print-administered (spiraled) samples of similar size. This results in more clustering, since students within a school tend to have more similar performance than students chosen in a simple random sample. Samples for the state assessment tend to have smaller design effects than the national samples. However, the properties of the state samples vary considerably.

The estimates for the total population tend to have the largest design effects, while the estimates for subpopulations, such as those based on parents' education level, tend to have smaller design effects. The parent education categorization forms more homogeneous collections of schools that have students with more similar backgrounds and performance thus reducing the variation in the types of schools (and students) included in the population estimates.

Table 10-11

Distributions of Design Effects by Demographic Subgroups for Proportion-correct Statistics Averaged Across Main Samples
Grade 4

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.42	1.63	1.84	2.61	1.61	0.39	33
Male	1.19	1.51	1.75	2.21	1.49	0.35	42
Female	1.26	1.43	1.62	2.41	1.47	0.35	40
White	1.36	1.57	1.76	2.46	1.54	0.36	32
Black	1.10	1.34	1.52	2.29	1.36	0.30	31
Hispanic	1.09	1.25	1.46	2.01	1.29	0.28	39
Asian American	0.84	1.03	1.18	2.01	1.06	0.33	23
Other Race/ethnicity	1.08	1.25	1.46	2.27	1.28	0.29	27
Other Type of Comm.	1.40	1.66	1.89	2.84	1.62	0.42	28
Disadvantaged Urban	1.15	1.38	1.71	2.79	1.43	0.43	21
Advantaged Urban	1.16	1.42	1.84	3.70	1.55	0.59	13
Par. Ed. < HS	1.10	1.28	1.62	2.44	1.36	0.38	36
Par. Ed. = HS	1.06	1.30	1.60	2.34	1.34	0.38	31
Par. Ed. > HS	1.13	1.30	1.52	2.19	1.34	0.31	32
Par. Ed. = College	1.19	1.37	1.61	2.27	1.41	0.31	36
Par. Ed. = IDK	1.22	1.44	1.69	2.34	1.46	0.33	30
Public School	1.40	1.56	1.76	2.45	1.55	0.35	34
Nonpublic School	1.08	1.25	1.57	2.71	1.34	0.41	20

Table 10-12

Distributions of Design Effects by Demographic Subgroups for Proportion-correct Statistics Averaged Across Main Samples

Grade 8

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.13	1.31	1.63	2.37	1.39	0.35	31
Male	1.04	1.29	1.48	2.19	1.28	0.30	35
Female	1.13	1.29	1.56	2.23	1.35	0.30	31
White	1.09	1.27	1.56	2.68	1.35	0.38	25
Black	1.06	1.23	1.48	2.34	1.28	0.33	26
Hispanic	0.97	1.14	1.46	2.54	1.23	0.36	24
Asian American	1.05	1.27	1.58	2.88	1.35	0.44	17
Other Race/ethnicity	0.92	1.11	1.35	2.50	1.16	0.36	16
Other Type of Comm.	1.06	1.32	1.62	2.78	1.39	0.40	25
Disadvantaged Urban	0.98	1.32	1.62	2.43	1.32	0.42	18
Advantaged Urban	1.08	1.38	1.88	2.90	1.51	0.53	13
Par. Ed. < HS	1.02	1.16	1.44	2.56	1.24	0.35	21
Par. Ed. = HS	1.01	1.25	1.51	2.31	1.27	0.32	31
Par. Ed. > HS	0.98	1.19	1.46	2.22	1.24	0.33	26
Par. Ed. = College	1.10	1.28	1.54	2.35	1.35	0.33	35
Par. Ed. = IDK	1.03	1.30	1.51	2.39	1.30	0.34	32
Public School	1.10	1.28	1.52	2.31	1.33	0.32	30
Nonpublic School	1.05	1.25	1.62	4.11	1.50	0.79	24

Table 10-13

Distributions of Design Effects by Demographic Subgroups for Proportion-correct Statistics Averaged Across Main Samples
Grade 12

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.28	1.51	1.69	2.61	1.52	0.34	32
Male	1.19	1.37	1.59	2.26	1.42	0.31	37
Female	1.20	1.38	1.64	2.66	1.43	0.36	36
White	1.27	1.49	1.77	2.58	1.52	0.35	31
Black	1.18	1.37	1.63	2.82	1.41	0.33	26
Hispanic	1.06	1.33	1.78	3.40	1.49	0.62	12
Asian American	1.14	1.39	1.56	3.05	1.41	0.40	20
Other Race/ethnicity	0.91	1.23	1.52	3.29	1.29	0.53	11
Other Type of Comm.	1.23	1.47	1.76	2.76	1.54	0.39	28
Disadvantaged Urban	1.07	1.30	1.58	2.71	1.34	0.40	21
Advantaged Urban	1.14	1.40	1.78	3.54	1.52	0.58	16
Par. Ed. < HS	1.07	1.22	1.42	2.18	1.26	0.30	33
Par. Ed. = HS	1.09	1.26	1.50	2.10	1.30	0.32	39
Par. Ed. > HS	1.09	1.26	1.52	2.26	1.34	0.36	33
Par. Ed. = College	1.18	1.41	1.64	2.29	1.42	0.34	40
Par. Ed. = IDK	1.05	1.39	1.61	2.88	1.41	0.51	26
Public School	1.22	1.38	1.59	2.50	1.43	0.31	32
Nonpublic School	1.16	1.45	2.05	3.87	1.65	0.70	19

Table 10-14

Distributions of Design Effects by Demographic Subgroups for Proportion-correct Statistics Averaged Across Long-Term Trend Reading and Writing Samples
Age 9

Subgroup	Quartile	Median	Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.13	1.31	1.42	1.80	1.29	0.22	33
Male	0.99	1.15	1.29	1.77	1.17	0.21	30
Female	1.07	1.22	1.58	2.26	1.33	0.33	25
White	1.08	1.18	1.42	1.78	1.23	0.25	30
Black	0.91	1.26	1.48	1.88	1.22	0.35	22
Hispanic	1.06	1.16	1.41	2.07	1.21	0.35	22
Asian American	0.65	0.95	1.18	2.38	0.98	0.44	7
Other Race/ethnicity	0.88	0.98	1.11	1.84	1.02	0.28	30
Other Type of Comm.	1.05	1.30	1.45	1.85	1.28	0.27	29
Disadvantaged Urban	0.91	1.06	1.46	2.38	1.24	0.45	12
Advantaged Urban	1.04	1.15	1.41	2.08	1.18	0.37	16
Par. Ed. < HS	0.91	1.14	1.26	1.89	1.16	0.31	31
Par. Ed. = HS	0.95	1.34	1.56	1.91	1.27	0.37	25
Par. Ed. > HS	1.07	1.16	1.32	1.79	1.22	0.25	33
Par. Ed. = College	1.07	1.15	1.44	1.88	1.25	0.26	37
Par. Ed. = IDK	1.08	1.16	1.30	1.57	1.15	0.22	38
Public School	1.11	1.34	1.53	1.89	1.30	0.27	34
Nonpublic School	0.94	1.08	1.29	2.07	1.14	0.40	16

Table 10-15

Distributions of Design Effects by Demographic Subgroups
for Proportion-correct Statistics Averaged Across Long-Term Trend Reading and Writing Samples
Age 13

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.21	1.37	1.75	2.46	1.46	0.39	25
Male	0.94	1.21	1.60	2.14	1.27	0.41	25
Female	1.09	1.20	1.47	1.87	1.27	0.26	29
White	1.11	1.32	1.55	2.29	1.37	0.31	20
Black	1.18	1.39	1.59	2.19	1.38	0.38	20
Hispanic	0.91	1.11	1.41	2.34	1.20	0.38	15
Asian American	0.88	1.19	1.55	3.36	1.30	0.62	5
Other Race/ethnicity	0.72	1.04	1.25	1.81	1.02	0.35	12
Other Type of Comm.	1.22	1.45	1.74	2.44	1.46	0.42	21
Disadvantaged Urban	0.91	1.28	1.59	2.21	1.28	0.41	16
Advantaged Urban	0.91	1.76	2.83	5.06	1.95	1.18	5
Par. Ed. < HS	0.77	1.02	1.23	1.62	1.01	0.28	27
Par. Ed. = HS	1.09	1.19	1.43	2.22	1.27	0.29	26
Par. Ed. > HS	1.02	1.26	1.49	1.97	1.28	0.31	30
Par. Ed. = College	1.02	1.39	1.61	2.20	1.34	0.36	31
Par. Ed. = IDK	1.07	1.24	1.48	2.08	1.25	0.28	25
Public School	1.09	1.28	1.67	2.28	1.34	0.36	24
Nonpublic School	1.24	1.51	1.76	3.29	1.64	0.61	15

Table 10-16

Distributions of Design Effects by Demographic Subgroups
for Proportion-correct Statistics Averaged Across Long-Term Trend Reading and Writing Samples
Age 17

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.19	1.45	1.71	2.21	1.45	0.31	31
Male	1.22	1.40	1.58	1.93	1.37	0.29	35
Female	1.11	1.26	1.45	1.83	1.28	0.24	30
White	1.22	1.46	1.65	2.16	1.49	0.30	26
Black	1.04	1.19	1.50	1.87	1.25	0.28	30
Hispanic	0.74	0.95	1.26	1.63	1.00	0.31	19
Asian American	0.86	1.11	1.63	2.83	1.29	0.56	14
Other Race/ethnicity	0.95	1.10	1.34	1.89	1.14	0.34	19
Other Type of Comm.	1.29	1.50	1.75	2.27	1.53	0.31	26
Disadvantaged Urban	0.94	1.24	1.66	2.12	1.29	0.41	24
Advantaged Urban	0.88	1.20	1.52	2.21	1.22	0.38	17
Par. Ed. < HS	0.93	1.11	1.25	1.70	1.10	0.26	21
Par. Ed. = HS	1.09	1.24	1.43	1.91	1.26	0.26	29
Par. Ed. > HS	1.16	1.35	1.61	2.14	1.41	0.33	28
Par. Ed. = College	1.09	1.36	1.62	2.08	1.38	0.31	30
Par. Ed. = IDK	0.80	0.99	1.37	1.82	1.07	0.37	29
Public School	1.13	1.47	1.64	2.17	1.41	0.34	34
Nonpublic School	1.17	1.58	1.97	2.60	1.61	0.49	16

Table 10-17

Distributions of Design Effects by Demographic Subgroups
for Proportion-correct Statistics Averaged Across Long-Term Trend Mathematics and Science Samples
Age 9

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.36	1.72	2.22	4.07	1.84	0.65	16
Male	1.18	1.36	1.68	2.93	1.47	0.43	23
Female	1.20	1.48	1.73	3.47	1.56	0.52	19
White	1.21	1.45	1.91	3.26	1.61	0.54	18
Black	0.93	1.19	1.54	2.80	1.28	0.47	15
Hispanic	1.01	1.28	1.57	2.72	1.34	0.43	19
Asian American	0.86	1.14	1.51	3.21	1.27	0.52	13
Other Race/ethnicity	0.95	1.10	1.38	2.68	1.22	0.42	20
Other Type of Comm.	1.36	1.73	2.32	4.08	1.86	0.70	15
Disadvantaged Urban	0.97	1.36	1.89	4.03	1.48	0.75	8
Advantaged Urban	0.98	1.47	1.99	3.54	1.53	0.72	11
Par. Ed. < HS	0.91	1.14	1.37	2.20	1.18	0.34	25
Par. Ed. = HS	0.97	1.16	1.43	3.15	1.22	0.41	19
Par. Ed. > HS	0.93	1.11	1.27	1.90	1.12	0.29	29
Par. Ed. = College	1.05	1.28	1.46	2.52	1.29	0.35	27
Par. Ed. = IDK	1.16	1.41	1.71	2.91	1.49	0.44	23
Public School	1.36	1.82	2.26	3.86	1.86	0.69	16
Nonpublic School	1.10	1.49	2.06	4.25	1.68	0.77	10

Table 10-18

Distributions of Design Effects by Demographic Subgroups
for Proportion-correct Statistics Averaged Across Long-Term Trend Mathematics and Science Samples
Age 13

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.32	1.63	2.06	3.51	1.72	0.55	22
Male	1.13	1.35	1.69	3.20	1.43	0.44	22
Female	1.15	1.43	1.78	2.52	1.47	0.43	25
White	1.24	1.53	1.84	3.56	1.59	0.51	21
Black	1.03	1.30	1.62	3.30	1.36	0.47	17
Hispanic	0.86	1.01	1.31	2.16	1.09	0.37	18
Asian American	0.76	1.02	1.31	2.91	1.09	0.47	11
Other Race/ethnicity	0.74	0.97	1.26	2.68	1.05	0.41	13
Other Type of Comm.	1.23	1.53	1.97	3.15	1.62	0.51	22
Disadvantaged Urban	1.00	1.55	2.23	6.24	1.75	1.03	6
Advantaged Urban	1.15	1.51	2.26	4.51	1.76	0.83	9
Par. Ed. < HS	0.84	1.04	1.30	2.07	1.09	0.33	22
Par. Ed. = HS	0.99	1.23	1.47	2.55	1.26	0.37	24
Par. Ed. > HS	0.94	1.13	1.39	2.08	1.18	0.33	26
Par. Ed. = College	1.12	1.39	1.69	2.33	1.43	0.39	28
Par. Ed. = IDK	0.94	1.13	1.38	2.03	1.16	0.30	29
Public School	1.29	1.59	2.03	3.22	1.68	0.53	21
Nonpublic School	1.33	1.76	2.32	4.39	1.92	0.82	12

Table 10-19

Distributions of Design Effects by Demographic Subgroups
for Proportion-correct Statistics Averaged Across Long-Term Trend Mathematics and Science Samples
Age 17

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.39	1.78	2.14	3.52	1.82	0.56	23
Male	1.28	1.58	1.92	3.17	1.64	0.47	27
Female	1.23	1.52	1.78	2.64	1.53	0.43	26
White	1.28	1.56	1.88	3.17	1.60	0.50	22
Black	1.10	1.45	1.90	4.19	1.55	0.65	12
Hispanic	1.00	1.20	1.61	3.83	1.39	0.63	10
Asian American	0.88	1.11	1.35	2.35	1.14	0.37	20
Other Race/ethnicity	0.86	1.06	1.35	2.82	1.14	0.44	13
Other Type of Comm.	1.38	1.74	2.30	4.07	1.86	0.68	16
Disadvantaged Urban	0.95	1.26	1.72	3.24	1.39	0.63	10
Advantaged Urban	1.18	1.63	2.08	4.13	1.72	0.73	11
Par. Ed. < HS	1.04	1.38	1.77	3.08	1.43	0.53	15
Par. Ed. = HS	1.15	1.42	1.67	2.25	1.42	0.36	32
Par. Ed. > HS	1.16	1.39	1.80	3.35	1.47	0.45	22
Par. Ed. = College	1.13	1.35	1.62	3.09	1.41	0.41	24
Par. Ed. = IDK	0.86	1.07	1.28	1.86	1.10	0.29	27
Public School	1.38	1.68	2.10	3.49	1.78	0.54	23
Nonpublic School	1.26	1.66	2.38	6.18	1.93	1.05	7

Table 10-20

Distributions of Design Effects by Demographic Subgroups for Proportion-correct Statistics Averaged Across State Samples
Mathematics - Grade 4

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	0.98	1.14	1.33	2.03	1.17	0.26	42
Male	0.94	1.09	1.26	1.96	1.11	0.24	44
Female	0.94	1.09	1.25	1.98	1.11	0.24	45
White	0.95	1.11	1.30	2.23	1.14	0.28	38
Black	0.85	1.05	1.28	2.34	1.09	0.36	24
Hispanic	0.89	1.04	1.21	2.09	1.07	0.27	35
Asian American	0.85	1.08	1.37	3.00	1.17	0.48	18
Other Race/ethnicity	0.86	1.04	1.28	2.78	1.10	0.37	25
Other Type of Comm.	0.97	1.13	1.32	2.03	1.16	0.27	38
Disadvantaged Urban	0.67	0.96	1.40	3.91	1.12	0.66	9
Advantaged Urban	0.61	0.94	1.43	5.27	1.13	0.80	7
Par. Ed. < HS	0.91	1.03	1.19	2.10	1.06	0.25	40
Par. Ed. = HS	0.91	1.05	1.21	2.03	1.07	0.24	41
Par. Ed. > HS	0.90	1.04	1.19	1.97	1.06	0.24	43
Par. Ed. = College	0.93	1.09	1.27	2.02	1.12	0.26	40
Par. Ed. = IDK	0.93	1.07	1.24	1.90	1.10	0.24	45

Table 10-21

Distributions of Design Effects by Demographic Subgroups for Proportion-correct Statistics Averaged Across State Samples Mathematics - Grade 8

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	0.97	1.14	1.32	2.19	1.16	0.27	40
Male	0.94	1.10	1.27	2.16	1.12	0.26	40
Female	0.94	1.09	1.27	1.98	1.12	0.25	41
White	0.94	1.11	1.31	2.50	1.16	0.33	36
Black	0.85	1.06	1.33	2.68	1.13	0.41	21
Hispanic	0.87	1.03	1.21	2.33	1.07	0.30	31
Asian American	0.87	1.10	1.42	3.29	1.20	0.53	17
Other Race/ethnicity	0.86	1.11	1.41	3.73	1.23	0.57	17
Other Type of Comm.	0.96	1.13	1.33	2.30	1.16	0.29	36
Disadvantaged Urban	0.66	0.99	1.43	3.73	1.12	0.64	8
Advantaged Urban	0.65	1.33	2.60	8.82	1.93	1.77	4
Par. Ed. < HS	0.89	1.04	1.19	2.09	1.06	0.26	37
Par. Ed. = HS	0.93	1.08	1.25	2.10	1.10	0.25	40
Par. Ed. > HS	0.92	1.06	1.23	2.20	1.09	0.26	41
Par. Ed. = College	0.94	1.10	1.27	2.08	1.12	0.26	40
Par. Ed. = IDK	0.91	1.04	1.20	2.05	1.07	0.25	40

Table 10-22

Distributions of Design Effects by Demographic Subgroups
for Proportion-correct Statistics Averaged Across State Samples
Reading - Grade 4

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	0.97	1.14	1.32	2.19	1.16	0.27	40
Male	0.94	1.10	1.27	2.16	1.12	0.26	40
Female	0.94	1.09	1.27	1.98	1.12	0.25	41
White	0.94	1.11	1.31	2.50	1.16	0.33	36
Black	0.85	1.06	1.33	2.68	1.13	0.41	21
Hispanic	0.87	1.03	1.21	2.33	1.07	0.30	31
Asian American	0.87	1.10	1.42	3.29	1.20	0.53	17
Other Race/ethnicity	0.86	1.11	1.41	3.73	1.23	0.57	17
Other Type of Comm.	0.96	1.13	1.33	2.30	1.16	0.29	36
Disadvantaged Urban	0.66	0.99	1.43	3.73	1.12	0.64	8
Advantaged Urban	0.65	1.33	2.60	8.82	1.93	1.77	4
Par. Ed. < HS	0.89	1.04	1.19	2.09	1.06	0.26	37
Par. Ed. = HS	0.93	1.08	1.25	2.10	1.10	0.25	40
Par. Ed. > HS	0.92	1.06	1.23	2.20	1.09	0.26	41
Par. Ed. = College	0.94	1.10	1.27	2.08	1.12	0.26	40
Par. Ed. = IDK	0.91	1.04	1.20	2.05	1.07	0.25	40

The tables show that the design effects are predominantly larger than 1, indicating that standard variance estimation formulas will be generally too small, sometimes markedly so. Although the distributions of design effects appear somewhat different for certain subgroups of the population, they are, perhaps, similar enough (at least within a grade) to select an overall composite value that is adequate for most purposes. In choosing a composite design effect, some consideration must be given to the relative consequences of overestimating the variance as opposed to underestimating the variance. For example, adopting the position that an overestimate of the variance is as severe an error as an underestimate leads to using a composite that is near to the center of the distributions of the design effects. Possible composites of this type are the mean and median design effects across the combined distribution of all design effects. In the current data, the mean design effects for total population estimates (which tend to be larger than most subpopulation estimates), respectively for age classes 9, 13, and 17, are 1.61, 1.39, and 1.52 for main samples, 1.29, 1.46, and 1.45 for reading and writing long-term trend samples, and 1.84, 1.72, and 1.82 for mathematics and science long-term trend samples. These are close to the median design effects: 1.63, 1.31, and 1.51 for main samples, 1.31, 1.37, and 1.45 for reading and writing long-term trend samples, and 1.72, 1.63, and 1.78 for mathematics and science long-term trend samples.

Alternatively, one can adopt the position that it is a graver error to underestimate the variability of a statistic than to overestimate it. For example, Johnson and King (1987) examine estimation of variances using design effects (among other techniques) under the assumption that the consequences of an underestimate are three times as severe as those of an overestimate of the same magnitude. Adopting a loss function that is a weighted sum of absolute values of the deviations of predicted from actual, with underestimates receiving three times the weight of overestimates, produces the upper quartile of the design effects as the composite value. This assumes that the size of the design effects do not depend on the size of the variance estimates. The values of this composite, respectively for age classes 9, 13, and 17, are 1.84, 1.63, and 1.69 for main samples, 1.42, 1.75, and 1.71 for reading and writing long-term trend samples, and 2.22, 2.06, and 2.14 for mathematics and science long-term trend samples.

10.4 THE DEGREES OF FREEDOM OF THE VARIANCE ESTIMATE

It is important to have an indication of the number of degrees of freedom to attribute to the jackknife variance estimator $\hat{Var}(t)$. The degrees of freedom of a variance estimator provide information on the stability of that estimator: the higher the number of degrees of freedom, the lower the variability of the estimator. In practical terms, the number of degrees of freedom of the variance estimator corresponds to the number of residual degrees of freedom that can be assumed for inferential procedures.

Since the jackknife procedure estimates the sampling variability of the statistic by assessing the effect of change in the sample at the paired first-stage sampling unit (FSSU) level, the number of degrees of freedom of the variance estimator $\hat{Var}(t)$ will be at most equal to M, the number of FSSU pairs. The maximum number of degrees of freedom equals the number of independent pieces of information used to generate the variance. In the case of data from the main assessments, the pieces of information are the 56 squared differences $(t_i - t)^2$, each supplying at most one degree of freedom (regardless of how many individuals were sampled within any FSSU).

The number of degrees of freedom of the sample variance estimator can be strictly less than the number of FSSU pairs. For example, suppose that the statistic t is a mean for some subgroup and no members of that subgroup can come from either FSSU in the ith FSSU pair. (Examples of such a subgroup are any PSU-level partitioning of the population, such as region.) In this instance, neither member of the FSSU pair i directly contributes to the estimate of t, so that the pseudoreplicate t_i would nearly equal the statistic t. If the replicate weights used to generate t_i had not received poststratification adjustments, the resulting pseudoreplicate t_i would be identical to the overall estimate t so that $(t_i - t)^2 = 0$. In this case, such a FSSU pair would impart no information to the variability of the statistic t and thus contribute zero degrees of freedom to the variance. However, since the replicate weights have received poststratification adjustments, the component $(t_i - t)^2$ is measuring the effect of the poststratification on the estimate. While being nonzero, such a component will tend to be much smaller in magnitude than the squared difference $(t_k - t)^2$ for any PSU pair k that does contribute to the estimate of t (see Rust, 1985).

The squared difference $(t_i - t)^2$ estimates σ_i^2, say, the contribution to the sampling variance of the statistic t which can be attributed to the ith FSSU pair and $\hat{Var}(t)$ estimates the sum of the contributions across all pairs:

$$\sum_{i=1}^{M} \sigma_i^2 \ .$$

If the σ_i^2 vary widely, as when a few of the σ_i^2 are markedly larger than the remainder, as in the above case where neither member of an FSSU pair contributes to the estimate of t, then $\hat{Var}(t)$ is predominantly estimating the sum of these larger components, which dominate the remaining terms. The effective degrees of freedom of $\hat{Var}(t)$ in this case will be nearer to the number of dominant terms.

One way to estimate how many degrees of freedom to attribute to the jackknife variance estimate of a statistic t is to match estimates of the first two moments of $\hat{Var}(t)$ to those of a chi-square random variable (Satterthwaite, 1941). If the t_i are normally distributed, the effective number of degrees of freedom using this approximation is

$$df_{eff} = \frac{\left(\sum_{i=1}^{M} (t_i - t)^2\right)^2}{\sum_{i=1}^{M} (t_i - t)^4}$$

However, empirical evidence from simulations indicates that the above formula has a severe downward bias in the case of the sum of single degree of freedom chi-square random variables (Johnson & Rust, 1992). More direct ways of assessing the effective degrees of freedom of a variance estimate are possible when a number of independent replicates of the estimate are available.

It is possible to estimate the number of degrees of freedom to attribute to the jackknife variance estimates of the weighted proportion-correct statistics by considering the distribution of design effects for a given set of items in a population or subpopulations (such as males or total) under the assumptions that the individual design effects are all estimating the same, underlying, design effect D and that the variance estimates of all weighted proportion-correct statistics have the same degrees of freedom, f. Specifically, assume that the jackknife variance estimate, V_j, of the j^{th} weighted proportion-correct statistic, P_j, is distributed like the random variable $(\sigma_j^2 / f)\, X_f^2$, where X_f^2 is a chi-square random variable with f degrees of freedom and σ_j^2 is the expected value of V_j. Further assume that the expected value of the conventional variance estimate is σ_j^2/D, where D is the underlying design effect. Then, for a sufficiently large sample size, so that the conventional variance estimate can be taken to be σ_j^2/D, the design effect of P_j will be approximately distributed like the constant (D/f) times a chi-square random variable with f degrees of freedom. If the underlying design effect D and the degrees of freedom f are the same for all P_j, then the distribution of the estimated design effects of the proportions correct across the set of items will be approximately distributed like a multiple times a chi-square random variable with f effective degrees of freedom.

An estimate effective degrees of freedom comes by matching moments of the design effects $D_1, \ldots, D_I$ across I items:

$$\bar{D} = E\left(\frac{D}{f}\chi_f^2\right) = \frac{D}{f} f = D$$

$$\hat{\sigma}_D^2 = \mathrm{var}\left(\frac{D}{f}\chi_f^2\right) = \frac{D^2}{f^2}(2f) = \frac{2}{f}D^2$$

which implies

$$\hat{\sigma}_D^2 = 2\,\frac{(\bar{D})^2}{f}$$

or

$$\hat{f} = 2\left(\frac{\bar{D}}{\hat{\sigma}_D}\right)^2 .$$

The tables in Appendix D show the result of this estimation of the effective degrees of freedom of the design effects, and hence the jackknife variance estimates, for weighted proportion-correct statistics for all samples in the 1992 assessment. The effective degrees of freedom in these tables are summarized in Tables 10-11 through 10-22.

The numbers in the tables show that the effective degrees of freedom of the jackknife variance estimates are indeed no larger than the number of FSSU pairs, and are, in fact, markedly smaller in some cases. The Asian American population is a example of a subgroup that has consistently small estimates of degrees of freedom. This is due to the fact that this population is concentrated in relatively few of the primary sampling units.

The effective degrees of freedom for the NAEP jackknife variance estimates are much smaller than the degrees of freedom attributed to the corresponding error estimates from conventional techniques. This fact affects inferential procedures since significance tests based on the conventional degrees of freedom will be too liberal (and confidence intervals will be too small). Fortunately, for the usual significance levels, the effect of using the effective degrees of freedom rather than the conventional values is generally moderate: a t statistic significant at the $\alpha = 5\%$ level assuming infinite degrees of freedom (essentially the conventional estimate) is significant at the $\alpha = 6\%$ level for 20 effective degrees of freedom, the $\alpha = 7\%$ level for 10 effective degrees of freedom, and the $\alpha = 10\%$ level for five effective degrees of freedom.

For practical purposes, the impact of the reduced degrees of freedom on inferential techniques can be largely accounted for by (1) using a moderate number (say 25) of degrees of freedom for all inferences about subgroups that appear approximately uniformly in all PSUs, and (2) using a smaller number (say 10) for the remaining subgroups. Certainly one should be cautious about barely significant results for subgroups that are highly clustered in the population.

Chapter 11

SCALING PROCEDURES

Eugene G. Johnson, Robert J. Mislevy, and Neal Thomas

Educational Testing Service

11.1 OVERVIEW

The primary method by which results from the 1992 National Assessment of Educational Progress were disseminated was scale-score reporting. With scaling methods, the performance of a sample of students in a subject area or subarea can be summarized on a single scale or series of scales even when different students have been administered different items. This chapter presents an overview of the scaling methodologies employed in the analyses of the data from NAEP surveys in general and from the 1992 assessment in particular. Details of the scaling procedures specific to the subject areas of reading, mathematics, science, and writing are presented in Chapters 12, 13, 14, and 15.

11.2 BACKGROUND

The basic information from an assessment consists of the responses of students to the items presented in the assessment. For NAEP, these items are generated to measure performance on sets of objectives developed by nationally representative panels of learning area specialists, educators, and concerned citizens. Satisfying the objectives of the assessment and ensuring that the tasks selected to measure each goal cover a range of difficulty levels typically require many items (e.g., the mathematics assessment required 205 items at grade 8. Depending on the subject areas, a mixture of multiple-choice, short constructed-response, and extended constructed-response items were used. Multiple-choice and short constructed-response items were used in all assessments but writing. Extended constructed-response items, scored on a multipoint scale, were presented in the main reading, mathematics, and writing assessments and in the long-term trend writing assessment. To reduce student burden, each assessed student was presented only a fraction of the full pool of items through multiple matrix sampling procedures.

The most direct manner of presenting the assessment results is to report separate statistics for each item. However, because of the vast amount of information, having separate results for each of the items in the assessment pool hinders the comparison of the general performance of subgroups of the population. Item-by-item reporting ignores overarching similarities in trends and subgroup comparisons that are common across items.

An obvious summary of performance across a collection of items is the average of the separate item scores. The advantage of averaging is that it tends to cancel out the effects of peculiarities in items that can affect item difficulty in unpredictable ways. Furthermore, averaging makes it possible to compare more easily the general performances of subpopulations.

Despite their advantages, there are a number of significant problems with average item scores. First, the interpretation of these results depends on the selection of the items; the selection of easy or difficult items could make student performance appear to be overly high or low. Second, the average score is related to the particular items comprising the average, so that direct comparisons in performance between subpopulations require that those subpopulations have been administered the same set of items. Third, because this approach limits comparisons to average scores on specific sets of items, it provides no simple way to report trends over time when the item pool changes. Finally, direct estimates of statistics such as the proportion of students who would achieve a certain score across the items in the pool are not possible when every student is administered only a fraction of the item pool. While the average score across all items in the pool can be readily obtained (as the average of the individual item scores), distributional statistics, such as quantiles of the distribution of scores across the full set of items, cannot be readily obtained without additional assumptions.

These limitations can be overcome by the use of response scaling methods. If several items require similar skills, the regularities observed in response patterns can often be exploited to characterize both respondents and items in terms of a relatively small number of variables. These variables include a respondent-specific variable, called proficiency, which quantifies a respondent tendency to answer items correctly (or, for multipoint items, to achieve a certain score) and item-specific variables that indicate characteristics of the item such as its difficulty, ability to distinguish between individuals with different levels of proficiency, and the chances of a very low proficiency respondent correctly answering a multiple-choice item. (These variables are discussed in more detail in the next section.) When combined through appropriate mathematical formulas, these variables capture the dominant features of the data. Furthermore, all students can be placed on a common scale, even though none of the respondents take all of the items within the pool. Using the scale, it becomes possible to discuss distributions of proficiency in a population or subpopulation and to estimate the relationships between proficiency and background variables.

It is important to point out that any procedure of aggregation, from a simple average to a complex multidimensional scaling model, highlights certain patterns at the expense of other potentially interesting patterns that may reside within the data. Every item in a NAEP survey is of interest and can provide useful information about what young Americans know and can do. The choice of an aggregation procedure must be driven by a conception of just which patterns are salient for a particular purpose.

The scaling for the assessments in reading and mathematics was carried out separately within the content areas specified in the framework for those subjects. This scaling within subareas was done because it was anticipated that different patterns of performance might exist for these essential subdivisions of the subject area. By creating a separate scale for each of these content areas, potential differences in subpopulation performance between the content areas are maintained. (Because of the small number of items within each of the purposes of writing, scaling was not carried out separately by purpose of writing.)

The creation of a series of separate scales to describe performance within a subject area does not preclude the reporting of an overall composite as a single index of overall performance in the subject area. A composite is computed as the weighted average of the content area scales, where the weights correspond to the relative importance given to each content area as defined by the framework. The composite provides a global measure of performance within the subject area, while the constituent content area scales allow the measurement of important interactions within educationally relevant subdivisions of the subject area.

11.3 SCALING METHODOLOGY

This section reviews the scaling models employed in the analyses of data from the 1992 assessment, and the multiple imputation or "plausible values" methodology that allows such models to be used with NAEP's sparse item-sampling design. The reader is referred to Mislevy (1991) for an introduction to plausible values methods and a comparison with standard psychometric analyses, to Mislevy, Johnson and Muraki (1992) and Beaton and Johnson (1992) for additional information on how the models are used in NAEP, and to Rubin (1987) for the theoretical underpinnings of the approach. It should be noted that the imputation procedure used by NAEP is a mechanism for providing plausible values for the unobserved proficiencies and not for filling in blank responses to background or cognitive variables.

While the NAEP procedures were developed explicitly to handle the characteristics of NAEP data, they build on other research, and are paralleled by other researchers. See, for example Dempster, Laird, and Rubin (1977); Little and Rubin (1983, 1987); Andersen (1980); Engelen (1987); Hoijtink (1991); Laird (1978); Lindsey, Clogg, and Grego (1991); Zwinderman (1991); Tanner and Wong (1987); and Rubin (1987, 1991).

11.3.1 The Scaling Models

Three distinct scaling models were used in the analysis of the data from the 1992 assessment. Each of the models is based on item response theory (IRT; e.g., Lord, 1980). Each is a "latent variable" model, defined separately for each of the scales, and quantifying respondents' tendencies to achieve a certain scores (such as correct/incorrect) on the items contributing to a scale as a function of a parameter that is not directly observed, called proficiency on the scale.

A three-parameter logistic (3PL) model was used for the multiple-choice items (which were scored correct/incorrect). The fundamental equation of the 3PL model is the probability that a person whose proficiency on scale k is characterized by the *unobservable* variable θ_k will respond correctly to item j:

$$P(X_j = 1|\theta_k, a_j, b_j, c_j) = c_j + \frac{(1 - c_j)}{1 + \exp[-1.7a_j(\theta_k - b_j)]} \qquad (11.1)$$

$$\equiv P_{j1}(\theta_k) \quad ,$$

where

x_j	is the response to item j, 1 if correct and 0 if not;
a_j	where $a_j > 0$, is the slope parameter of item j, characterizing its sensitivity to proficiency;
b_j	is the threshold parameter of item j, characterizing its difficulty; and
c_j	where $0 \le c_j < 1$, is the lower asymptote parameter of item j, reflecting the chances of students of very low proficiency selecting the correct option.

Further define the probability of an incorrect response to the item as

$$P_{j0} = P(x_j = 0 | \theta_k, a_j, b_j, c_j) = 1 - P_{j1}(\theta_k) \tag{11.2}$$

A two-parameter logistic (2PL) model was used for short constructed-response items, which were scored correct or incorrect. The form of the 2PL model is the same as equations (11.1) and (11.2) with the c_j parameter fixed at zero.

In addition to the multiple-choice and short constructed-response items, a number of extended constructed-response items were presented in the assessments of reading and mathematics; and only extended constructed-response items were presented in the writing assessments. Each of these items was scored on a multipoint scale with potential scores ranging from 0 to 4 for reading, mathematics and long-term trend writing and ranging from 0 to 6 for the main writing assessment. Additionally, as discussed in Chapter 13, in mathematics certain sets of items consisting of highly correlated parts were combined into "testlets" (Wainer & Kiely, 1987) where the score assigned to a testlet was the number of constituent parts answered correctly. Items that are scored on a multipoint scale are referred to as polytomous items, in contrast with the multiple-choice and short constructed-response items, which are scored correct/incorrect and referred to as dichotomous items.

The polytomous items were scaled using a generalized partial credit model (Muraki, 1992). The fundamental equation of this model is the probability that a person with proficiency θ_k on scale k will have, for the jth polytomous item, a response x_j that is scored in the ith of m_j ordered score categories:

$$P(X_j = i | \theta_k, a_j, b_j, d_{j,1}, \ldots, d_{j,m_j-1}) = \frac{\exp\left(\sum_{v=0}^{i} 1.7a_j(\theta_k - b_j + d_{j,v})\right)}{\sum_{g=0}^{m_j-1} \exp\left(\sum_{v=0}^{g} 1.7a_j(\theta_k - b_j + d_{j,v})\right)} \tag{11.3}$$

$$\equiv P_{ji}(\theta_k)$$

where

m_j	is the number of categories in the response to item *j*
x_j	is the response to item *j*, with possibilities $0,1,...,m_j-1$
a_j	is the slope parameter;
b_j	is the item location parameter characterizing overall difficulty; and
$d_{j,i}$	is the category *i* threshold parameter (see below).

Indeterminacies in the parameters of the above model are resolved by setting $d_{j,0} = 0$ and setting $\sum_{i=1}^{m_j-1} d_{j,i} = 0$. Muraki (1992) points out that $b_j - d_{j,i}$ is the point on the θ_k scale at which the plots of $P_{j,i-1}(\theta_k)$ and $P_{ji}(\theta_k)$ intersect and so characterizes the point on the θ_k scale at which the response to item *j* has the highest probability of incurring a change from response category *i*-1 to *i*.

When $m_j = 2$, so that there are two score categories (0,1), it can be shown that $P_{ji}(\theta_k)$ of equation 11.3 for $i=0,1$ corresponds respectively to $P_{j0}(\theta_k)$ and $P_{j1}(\theta_k)$ of the 2PL model (equations 11.1 and 11.2 with $c_j=0$).

For the purposes of reporting item parameter estimates and other intermediary estimates, the linear indeterminacies apparent in (11.1) and (11.3) may be resolved by an arbitrary choice of the origin and unit size in a given scale. For example, a provisional scale was employed in the analysis of the 1992 reading assessment by standardizing the combined age 9/grade 4, age 13/grade 8, and age 17/grade 12 samples. Final results for each content area were scale were linearly transformed from the θ scale to a 0-to-500 scale. Analogous scaling conventions and reporting transformations for the remaining 1992 assessments are described in the corresponding subject area chapters in this report.

A typical assumption of item response theory is the conditional independence of the response by an individual to a set of items, given the individual's proficiency. That is, conditional on the individual's θ_k, the joint probability of a particular response pattern $\underset{\sim}{x} = (x_1,...,x_n)$ across a set of *n* items is simply the product of terms based on (11.1), (11.2), and (11.3):

$$P(\underset{\sim}{x}|\theta_k, \textit{item parameters}) = \prod_{j=1}^{n} \prod_{i=0}^{m_j-1} P_{ji}(\theta_k)^{u_{ji}} \tag{11.4}$$

where $P_{ji}(\theta_k)$ is of the form appropriate to the type of item (dichotomous or polytomous), m_j is taken equal to 2 for the dichotomously scored items, and u_{ji} is an indicator variable defined by

$$u_{ji} = \begin{cases} 1 \text{ if response } x_j \text{ was in category } i \\ 0 \text{ otherwise.} \end{cases}$$

It is also typically assumed that response probabilities are conditionally independent of background variables ($\underline{y}$), given θ_k, or

$$P(\underline{x}|\theta_k, \textit{item parameters}, \underline{y}) = p(\underline{x}|\theta_k, \textit{item parameters}). \tag{11.5}$$

After $\underline{x}$ has been observed, equation 11.4 can be viewed as a likelihood function, and provides a basis for inference about θ_k or about item parameters. Estimates of item parameters were obtained by the NAEP BILOG/PARSCALE program, which combines Mislevy and Bock's (1982) BILOG and Muraki and Bock's (1991) PARSCALE computer programs, and which concurrently estimates parameters for all items (dichotomous and polytomous). The item parameters are then treated as known in subsequent calculations. In subject areas with multiple scales (reading and mathematics), the parameters of the items constituting each of the separate scales were estimated independently of the parameters of the other scales. Once items have been calibrated in this manner, a likelihood function for the scale proficiency θ_k is induced by a vector of responses to any subset of calibrated items, thus allowing θ_k-based inferences from matrix samples.

In all NAEP IRT analyses, missing responses at the end of each block a student was administered were considered "not-reached," and treated as if they had not been presented to the respondent. Missing responses to dichotomous items before the last observed response in a block were considered intentional omissions, and treated as fractionally correct at the value of the reciprocal of the number of response alternatives. These conventions are discussed by Mislevy and Wu (1988). With regard to the handling of not-reached items, Mislevy and Wu found that ignoring not-reached items introduces slight biases into item parameter estimation to the degree that not-reached items are present and speed is correlated with ability. With regard to omissions, they found that the method described above provides consistent limited-information likelihood estimates of item and ability parameters under the assumption that respondents omit only if they can do no better than responding randomly.

In mathematics, the extended constructed-response items were always the last item in a block. Because of this and because considerably more effort was required of the student to answer these items, nonresponse to an extended constructed-response item was considered an intentional omission (and scored as the lowest category, 0) unless the student also did not respond to the item immediately preceding that item. In that case, the extended constructed-response item was considered not reached and treated as if it had not been presented to the student.

Although the IRT models are employed in NAEP only to summarize performance, a number of checks are made to detect serious violations of the assumptions underlying the models (such as conditional independence). When warranted, remedial efforts are made to mitigate the effects of such violations on inferences. These checks include comparisons of empirical and theoretical item response functions to identify items for which the IRT model may provide a poor fit to the data.

Scaling areas in NAEP are determined *a priori* by considerations of content as collections of items for which overall performance is deemed to be of interest, as defined by the frameworks developed by the National Assessment Governing Board. A proficiency scale θ_k is defined *a priori* by the collection of items representing that scale. What is important, therefore, is that the models capture salient variation in the response data to effectively summarize the overall performance on the content area of the populations and subpopulations being assessed. Because of the *a priori* definition of the latent proficiency variable, departure from conditional independence tends to cancel out over items and does not seriously affect the estimation of whole group and subpopulation distributions, except when substantial differential item functioning (DIF) is found simultaneously for many items. NAEP has routinely conducted DIF analyses to guard against potential biases in making subpopulation comparisons based on the proficiency distributions.

The local independence assumption embodied in equation 11.4 implies that item response probabilities depend only on θ and the specified item parameters, and not on the position of the item in the booklet, the content of items around an item of interest, or the test-administration timing conditions. However, these effects are certainly present in any application. The practical question is whether inferences based on the IRT probabilities obtained via 11.4 are robust with respect to the ideal assumptions underlying the IRT model. Our experience with the 1986 NAEP reading anomaly (Beaton & Zwick, 1990) has shown that for measuring small changes over time, changes in item context and speededness conditions can lead to unacceptably large random error components. These can be avoided by presenting items used to measure change in identical test forms, with identical timings and administration conditions. Thus, we do *not* maintain that the item parameter estimates obtained in any particular booklet configuration are appropriate for other conceivable configurations. Rather, we assume that the parameter estimates are context-bound. (For this reason, we prefer common population equating to common item equating whenever equivalent random samples are available for linking.) This is the reason that the data from the Trial State Assessment were calibrated separately from the data from the national NAEP—since the administration procedures differed somewhat between the Trial State Assessment and the national NAEP, the values of the item parameters could be different.

11.3.2 An Overview of Plausible Values Methodology

Item response theory was developed in the context of measuring individual examinees' abilities. In that setting, each individual is administered enough items (often 60 or more) to permit precise estimation of his or her θ, as a maximum likelihood estimate $\hat{\theta}$, for example. Because the uncertainty associated with each θ is negligible, the distribution of θ, or the joint distribution of θ with other variables, can then be approximated using individuals' $\hat{\theta}$ values as if they were θ values.

This approach breaks down in the assessment setting when, in order to provide broader content coverage in limited testing time, each respondent is administered relatively few items in a scaling area. The problem is that the uncertainty associated with individual θs is too large to ignore, and the features of the $\hat{\theta}$ distribution can be seriously biased as estimates of the θ distribution. (The failure of this approach was verified in early analyses of the 1984 NAEP reading survey; see Wingersky, Kaplan, & Beaton, 1987.) Plausible values were developed as a way to estimate key population features consistently, and approximate others no worse than standard IRT procedures would. A detailed development of plausible values methodology is given in Mislevy (1991). Along with theoretical justifications, that paper presents comparisons with standard procedures, discussions of biases that arise in some secondary analyses, and numerical examples.

The following provides a brief overview of the plausible values approach, focusing on its implementation in the 1992 NAEP analyses.

Let $\underline{y}$ represent the responses of all sampled examinees to background and attitude questions, along with design variables such as school membership, and let $\underline{\theta}$ represent the subscale proficiency values. If $\underline{\theta}$ were known for all sampled examinees, it would be possible to compute a statistic $t(\underline{\theta},\underline{y})$—such as a scale or composite subpopulation sample mean, a sample percentile point, or a sample regression coefficient—to estimate a corresponding population quantity T. A function $U(\underline{\theta},\underline{y})$—e.g., a jackknife estimate—would be used to gauge sampling uncertainty, as the variance of t around T in repeated samples from the population.

Because the scaling models are latent variable models, however, $\underline{\theta}$ values are not observed even for sampled students. To overcome this problem, we follow Rubin (1987) by considering $\underline{\theta}$ as "missing data" and approximate $t(\underline{\theta},\underline{y})$ by its expectation given $(\underline{x},\underline{y})$, the data that actually were observed, as follows:

$$\begin{aligned} t^*(\underline{x},\underline{y}) &= E[t(\underline{\theta},\underline{y})\,|\,\underline{x},\underline{y}] \\ &= \int t(\underline{\theta},\underline{y})\, p(\underline{\theta}\,|\,\underline{x},\underline{y})\, d\underline{\theta}\,. \end{aligned} \tag{11.6}$$

It is possible to approximate t^* using random draws from the conditional distribution of the scale proficiencies given the item responses x_i, background variables y_i, and model parameters for sampled student i. These values are referred to as imputations in the sampling literature, and plausible values in NAEP. The value of $\underline{\theta}$ for any respondent that would enter into the computation of t is thus replaced by a randomly selected value from their conditional distribution. Rubin (1987) proposes that this process be carried out several times—"multiple imputations"—so that the uncertainty associated with imputation can be quantified. The average of the results of, for example, M estimates of t, each computed from a different set of plausible values, is a Monte Carlo approximation of (11.6); the variance among them, B, reflects uncertainty due to not observing θ, and must be added to the estimated expectation of $U(\underline{\theta},\underline{y})$, which reflects uncertainty due to testing only a sample of students from the population. Section 11.5 explains how plausible values are used in subsequent analyses.

It cannot be emphasized too strongly that **plausible values are *not* test scores for *individuals*** in the usual sense. Plausible values are offered only as intermediary computations for calculating integrals of the form of equation 11.6, in order to estimate *population* characteristics. When the underlying model is correctly specified, plausible values will provide consistent estimates of population characteristics, even though they are not generally unbiased estimates of the proficiencies of the individuals with whom they are associated. The key idea lies in a contrast between plausible values and the more familiar θ estimates of educational measurement that are in some sense optimal for each examinee (e.g., maximum likelihood estimates, which are consistent estimates of an examinee's θ, and Bayes estimates, which provide minimum mean-squared errors with respect to a reference population): *Point estimates that are optimal for individual examinees have distributions that can produce decidedly nonoptimal (specifically, inconsistent) estimates of population characteristics* (Little & Rubin, 1983). Plausible values, on the other hand, are constructed explicitly to provide consistent estimates of population effects. For further discussion see Mislevy, Beaton, Kaplan, and Sheehan (1992).

11.3.3 Computing Plausible Values in IRT-based Scales

Plausible values for each respondent *i* are drawn from the conditional distribution $p(\underline{\theta}_i|x_i,y_i,\Gamma,\Sigma)$, where Γ and Σ are regression model parameters defined in this subsection. This subsection describes how, in IRT-based scales, these conditional distributions are characterized, and how the draws are taken. An application of Bayes' theorem with the IRT assumption of conditional independence produces

$$p(\underline{\theta}_i|x_i,y_i,\Gamma,\Sigma) \propto P(x_i|\underline{\theta}_i,y_i,\Gamma,\Sigma)\, p(\underline{\theta}_i|y_i,\Gamma,\Sigma) = P(x_i|\underline{\theta}_i)\, p(\underline{\theta}_i|y_i,\Gamma,\Sigma) \ , \tag{11.7}$$

where, for vector-valued $\underline{\theta}_i$, $P(x_i|\underline{\theta}_i)$ is the product over scales of the *independent likelihoods* induced by responses to items within each scale, and $p(\underline{\theta}_i|y_i,\Gamma,\Sigma)$ is the multivariate—and generally nonindependent—*joint density* of proficiencies for the scales, conditional on the observed value y_i of background responses, and the parameters Γ and Σ. The scales are determined by the item parameter estimates that constrain the population mean to zero and standard deviation to one. The item parameter estimates are fixed and regarded as population values in the computation described in this subsection.

In the analyses of the data from the main assessments, a normal (Gaussian) form was assumed for $p(\underline{\theta}_i|y_i\Gamma,\Sigma)$, with a common variance, Σ, and with a mean given by a linear model with slope parameters, Γ, based on the first approximately principal components of several hundred selected main-effects and two-way interactions of the complete vector of background variables. The included principal components will be referred to as the *conditioning variables*, and will be denoted y^c. (The complete set of original background variables used in the analyses of each subject area are listed in Appendix F.) The following model was fit to the data within each subject area:

$$\theta = \Gamma' y^c + \varepsilon \ , \tag{11.8}$$

where ε is normally distributed with mean zero and variance Σ. The number of principal components of the conditioning variables used was sufficient to account for 90 percent of the total variance of the full set of conditioning variables (after standardizing each variable). As in regression analysis, Γ is a matrix each of whose columns is the *effects* for one scale and Σ is the matrix *variance of residuals* between scales.

A model similar to (11.8) was used for the long-term trend assessments, with the difference that y^c consisted of main effects and interactions from the smaller set of background variables available in the long-term trend assessments.

Maximum likelihood estimates of Γ and Σ, denoted by $\hat{\Gamma}$ and $\hat{\Sigma}$, are obtained from Sheehan's (1985) MGROUP computer program using the EM algorithm described in Mislevy (1985). The EM algorithm requires the computation of the mean, $\bar{\underline{\theta}}_i$, and variance, Σ_i^p, of the posterior distribution in (11.7). For subject areas with multiple scales, the CGROUP version of the MGROUP program was used to compute the moments using higher order asymptotic corrections to a normal approximation (Thomas, 1993). For writing and for the long-term trend assessments, each of which have a single scale, the BGROUP version of MGROUP was used. BGROUP uses numeric quadrature to evaluate the posterior moments required by the E-step of the EM algorithm for one- and two-dimensional applications (Thomas, 1993). For estimation of group means on a single scale, CGROUP and BGROUP results will be nearly identical to those from the original MGROUP program. CGROUP and BGROUP yield better estimates of correlations between scales, and hence better estimates of composite scale means. BGROUP will, theoretically, yield better estimates than CGROUP, but because of the methodology used, its function is limited to bivariate scales. Hence CGROUP is used for assessments involving more than two scales.

After completion of the EM algorithm, the plausible values are drawn in a three-step process from the joint distribution of the values of Γ for all sampled respondents. First, a value of Γ is drawn from a normal approximation to $P(\Gamma,\Sigma|x_i,y_i)$ that fixes Σ at the value $\hat{\Sigma}$, (Thomas, 1993). Second, conditional on the generated value of Γ (and the fixed value of $\Sigma = \hat{\Sigma}$), the mean, $\bar{\underline{\theta}}_i$, and variance, Σ_i^p, of the posterior distribution in equation 11.7 (i.e., $p(\underline{\theta}_i|x_i,y_i,\Gamma,\Sigma)$) are computed using the same methods applied in the EM algorithm. In the third step, the $\underline{\theta}_i$ are drawn independently from a multivariate normal distribution with mean $\bar{\underline{\theta}}_i$ and variance Σ_i^p, approximating the distribution in (8.7). These three steps are repeated five times producing five imputations of $\bar{\underline{\theta}}_i$ for each sampled respondent.

11.4 ACHIEVEMENT LEVELS AND SCALE ANCHORING

Since its beginning, a goal of NAEP has been to inform the public about what students in American schools know and can do. While the NAEP scales provide information about the distributions of proficiency for the various subpopulations, they do not directly provide information about the meaning of various points on the scale. Traditionally, meaning has been attached to educational scales by norm-referencing—that is, by comparing students at a particular scale level to other students. In contrast, NAEP achievement levels and scale anchors describe selected points on the scale in terms of the types of skills that are likely to be exhibited by students scoring at that level. Both the achievement level process and scale anchoring were applied to the mathematics composite. The achievement level process was applied to the reading composite and a modified anchoring process was also applied to the achievement level cutpoints.

The National Assessment Governing Board has determined that achievement levels shall be the first and primary way of reporting NAEP results. Setting achievement levels is a method for setting standards on the NAEP assessment that identify what students *should* know and be able to do at various points on the composite. For each grade and separately for mathematics and reading, three levels were defined—basic, proficient, and advanced. Based on initial policy definitions of these levels, panelists were asked to determine operational descriptions of the levels appropriate with the content and skills assessed in the reading assessment. With these descriptions in mind, the panelists were then asked to rate the assessment items in terms of the expected performance of marginally acceptable examinees at each of these three levels. These ratings were then mapped onto the NAEP scale to obtain the achievement level cutpoints for reporting. Further details of the achievement level-setting process appear in Appendix G for mathematics and Appendix H for reading.

The achievement level-setting process specifies expected performance of students at each of the three achievement levels. To determine the types of skills currently exhibited by students at each of the levels, ETS applied a modified anchoring procedure to the 1992 reading achievement levels. As applied to the achievement levels, the anchoring process was designed to determine the sets of questions that students scoring at or above each achievement level cutpoint could perform with a high degree of success. Specifically, a question was identified as anchoring at an achievement level for a given grade if it was answered correctly by at least 65 percent of the students in that grade scoring at the cutpoint of that achievement level, and by less than 65 percent of the students scoring at the cutpoints for any lower achievement level. A committee of reading experts, educators, and others was assembled to review the questions and, using their knowledge of reading and student performance, to generalize from the questions to descriptions of the types of skills exhibited at each achievement level. Further details of the anchoring process for reading appear in Appendix J.

As applied to the 1992 mathematics data, scale anchoring began by identifying four anchoring levels on the mathematics composite: 200, 250, 300, 350. The next step was to identify items that a large majority (at least 65 percent) of students at a given anchor level could answer correctly but that most students (at least 50 percent) at the next lower level answered incorrectly. Additionally, there had to be at least a 30 percentage point difference in the probabilities of success between the two levels. The result was a grouping of assessment items by the levels between which they discriminate. This is in contrast with the procedure used for reading, where the items were grouped by the first achievement level at which students had at least a 65 percent probability of success. Like the reading, the mathematics anchor items were then reviewed by subject area experts who, using their knowledge of mathematics and student performance, generalized from the items to descriptions of the types of skills exhibited at each level. Further details of the anchoring process for mathematics appear in Appendix I.

11.5 ANALYSES

When survey variables are observed without error from every respondent, standard variance estimators quantify the uncertainty associated with sample statistics from the only source of the uncertainty, namely the sampling of respondents. Item-level statistics for NAEP cognitive items meet this requirement, but scale-score proficiency values do not. The IRT models used in their construction posit an unobservable proficiency variable θ to summarize performance on the items in the subarea. The fact that θ values are not observed even for the respondents in the sample requires additional statistical analyses to draw inferences about θ distributions and to quantify the uncertainty associated with those inferences. As described above, Rubin's (1987) multiple imputations procedures were adapted to the context of latent variable models to produce the plausible values upon which many analyses of the data from the 1992 assessment were based. This section describes how plausible values were employed in subsequent analyses to yield inferences about population and subpopulation distributions of proficiencies.

11.5.1 Computational Procedures

Even though one does not observe the θ value of respondent i, one does observe variables that are related to it: x_i, the respondent's answers to the cognitive items he or she was administered in the area of interest, and y_i, the respondent's answers to demographic and background variables. Suppose one wishes to draw inferences about a number $T(\underline{\theta},\underline{Y})$ that could be calculated explicitly if the θ and y values of each member of the population were known. Suppose further that if θ values were observable, we would be able to estimate T from a sample of N pairs of θ and y values by the statistic $t(\underline{\theta},\underline{y})$ [where $(\underline{\theta},\underline{y}) \equiv (\theta_1,y_1,\ldots,\theta_N,y_N)$], and that we could estimate the variance in t around T due to sampling respondents by the function $U(\underline{\theta},\underline{y})$. Given that observations consist of (x_i,y_i) rather than $(\underline{\theta}_i,y_i)$, we can approximate t by its expected value conditional on $(\underline{x},\underline{y})$, or

$$t^{*}(\underline{x},\underline{y}) \;=\; E[t(\underline{\theta},\underline{y})\,|\,\underline{x},\underline{y}] \;=\; \int t(\underline{\theta},\underline{y})\, p(\underline{\theta}\,|\,\underline{x},\underline{y})\, d\underline{\theta}\,.$$

It is possible to approximate t^{*} with random draws from the conditional distributions $p(\underline{\theta}_i|x_i,y_i)$, which are obtained for all respondents by the method described in section 8.3.3. Let $\hat{\underline{\theta}}_m$ be the mth such vector of plausible values, consisting of a multidimensional value for the latent variable of each respondent. This vector is a plausible representation of what the true $\underline{\theta}$ vector might have been, had we been able to observe it.

The following steps describe how an estimate of a scalar statistic $t(\underline{\theta},\underline{y})$ and its sampling variance can be obtained from M (> 1) such sets of plausible values. (Five sets of plausible values are used in NAEP analyses.)

1) Using each set of plausible values $\hat{\underline{\theta}}_m$ in turn, evaluate t as if the plausible values were true values of $\underline{\theta}$. Denote the results $\hat{t}_m$, for $m=1,\ldots,M$.

2) Using the jackknife variance estimator defined in Chapter 10, compute the estimated sampling variance of $\hat{t}_m$, denoting the result U_m.

3) The final estimate of t is

$$t^{*} = \sum_{m=1}^{M} \frac{\hat{t}_m}{M}.$$

4) Compute the average sampling variance over the M sets of plausible values, to approximate uncertainty due to sampling respondents:

$$U^{*} = \sum_{m=1}^{M} \frac{U_m}{M}.$$

5) Compute the variance among the M estimates $\hat{t}_m$, to approximate uncertainty due to not observing θ values from respondents:

$$B = \sum_{m=1}^{M} \frac{(\hat{t}_m - t^*)^2}{(M - 1)}$$

6) The final estimate of the variance of t^* is the sum of two components:

$$V = U^* + (1 + M^{-1})\ B.$$

Note: Due to the excessive computation that would be required, NAEP analyses did not compute and average jackknife variances over all five sets of plausible values, but only on the first set. Thus, in NAEP reports, U^* is approximated by U_1.

11.5.2 Statistical Tests

Suppose that if θ values were observed for sampled students, the statistic $(t - T)/U^{1/2}$ would follow a t-distribution with d degrees of freedom. Then the incomplete-data statistic $(t^* - T)/V^{1/2}$ is approximately t-distributed, with degrees of freedom given by

$$\nu = \frac{1}{\dfrac{f^2}{M - 1} + \dfrac{(1 - f)^2}{d}}$$

where f is the proportion of total variance due to not observing θ values:

$$f = (1 + M^{-1})\ B/V.$$

When B is small relative to U^*, the reference distribution for incomplete-data statistics differs little from the reference distribution for the corresponding complete-data statistics. This is the case with main NAEP reporting variables. If, in addition, d is large, the normal approximation can be used to flag "significant" results.

For k-dimensional t, such as the k coefficients in a multiple regression analysis, each U_m and U^* is a covariance matrix, and B is an average of squares and cross-products rather than simply an average of squares. In this case, the quantity $(T-t^*)\ V^{-1}\ (T-t^*)'$ is approximately F distributed, with degrees of freedom equal to k and ν, with ν defined as above but with a matrix generalization of f:

$$f = (1 + M^{1})\ Trace\ (BV^{1})/k\ .$$

By the same reasoning as used for the normal approximation for scalar t, a chi-square distribution on k degrees of freedom often suffices.

11.5.3 Biases in Secondary Analyses

Statistics t^* that involve proficiencies in a scaled content area and variables included in the conditioning variables y^c are consistent estimates of the corresponding population values T. Statistics involving background variables y that were *not* conditioned on, or relationships among proficiencies from *different* content areas, are subject to asymptotic biases whose magnitudes depend on the type of statistic and the strength of the relationships of the nonconditioned background variables to the variables that were conditioned on and to the proficiency of interest. That is, the large sample expectations of certain sample statistics need not equal the true population parameters.

The *direction* of the bias is typically to underestimate the effect of nonconditioned variables. For details and derivations see Beaton and Johnson (1990), Mislevy (1991), and Mislevy and Sheehan (1987, section 10.3.5). For a given statistic t^* involving one content area and one or more nonconditioned background variables, the *magnitude* of the bias is related to the extent to which observed responses x account for the latent variable θ, and the degree to which the nonconditioned background variables are explained by conditioning background variables. The first factor—conceptually related to test reliability—acts consistently in that greater measurement precision reduces biases in *all* secondary analyses. The second factor acts to reduce biases in certain analyses but increase it in others. In particular,

- High shared variance between conditioned and nonconditioned background variables *mitigates* biases in analyses that involve only proficiency and nonconditioned variables, such as marginal means or regressions.
- High shared variance *exacerbates* biases in regression coefficients of conditional effects for nonconditioned variables, when nonconditioned and conditioned background variables are analyzed jointly as in multiple regression.

The large number of background variables that have been included in the conditioning vectors for the 1992 assessments allows a large number of secondary analyses to be carried out with little or no bias, and mitigates biases in analyses of the marginal distributions of θ in nonconditioned variables. Kaplan and Nelson's analysis of the 1988 NAEP reading data (some results of which are summarized in Mislevy, 1991), which had a similar design and fewer conditioning variables, indicate that the potential bias for nonconditioned variables in multiple regression analyses is below 10 percent, and biases in simple regression of such variables is below 5 percent. Additional research (summarized in Mislevy, 1990) indicates that most of the bias reduction obtainable from conditioning on a large number of variables can be captured by instead conditioning on the first several principal components of the matrix of all original conditioning variables. This procedure was adopted for the 1992 main assessments by replacing the conditioning effects by the first K principal components, where K was selected so that 90 percent of the total variance of the full set of conditioning variables (after standardization) was captured. Mislevy (1990) shows that this puts an upper bound of 10 percent on the average bias for all analyses involving the original conditioning variables.

11.5.4 A Numerical Example

To illustrate how plausible values are used in subsequent analyses, this subsection gives some of the steps in the calculation of the 1992 grade 4 reading composite mean and its estimation-error variance.

The weighted mean of the first plausible values of the reading composite for the grade 4 students in the sample is 217.79, and the jackknife variance of these values is 0.833. Were these values true θ values, then 217.79 would be the estimate of the mean and 0.833 would be the estimation-error variance. The weighted mean of the second plausible values of the same students, however, is 217.62; the third, fourth, and fifth plausible values give weighted means of 217.74, 218.24, and 218.05. Since all of these figures are based on precisely the same sample of students, the variation among them is due to uncertainty about the students' θs, having observed their item responses and background variables. Consequently, our best estimate of the mean for grade 4 students is the average of the five plausible values: 217.89. Taking the jackknife variance estimate from the first plausible value, 0.833, as our estimate U^* of sampling variance, and the variance among the five weighted means, .063, as our estimate B of uncertainty due to not observing θ, we obtain as the final estimate V of total error variance $0.833 + (1+5^{-1})\ .063 = 0.908$.

It is also possible to partition the estimation error variance of a statistic using these same variance components. The proportion of error variance due to sampling students from the population is U^*/V, and the proportion due to the latent nature of θ is $(1+M^{-1})B/V$. The results are shown in Table 11-1. The value of U^*/V roughly corresponds to reliability in classical test theory and indicates the amount of information about an average individual's θ present in the observed responses of the individual.

Table 11-1
Estimation Error Variance and Related Coefficients
for the 1992 Grade 4 Reading Composite
(Based on Five Plausible Values)

			Proportion of Variance Due to...	
U^*	**$(1+5^{-1})B$**	**V**	**Student Sampling: U^*/V**	**Latency of θ: $(1+5^{-1})B/V$**
0.833	0.076	0.908	0.92	0.08

Chapters 12, 13, 14, and 15 and Appendix N provide values of the proportion of variance due to sampling and due to the latent nature of θ for all 1992 scales and composites for the populations as a whole and, in the appendix, for selected subpopulations. It will be seen that the proportion of variance due to the latency of θ varies among subject area, tending to be small for mathematics, where there are many items and highly correlated scales, larger for reading, where the scales are less highly correlated, and largest for writing, where there is low correlation between tasks and each student responded to only one or two tasks. Essentially, the variance due to the latent nature of θ is largest when there is less information about a student's proficiency. Given fixed assessment time, this decrease in information will occur whenever the amount of information per unit time decreases as can happen when many short constructed-response or multiple-choice items are replaced by a few extended constructed-response items.

11.6 OVERVIEW OF THE 1992 NAEP SCALES

IRT scale-score analyses were carried out in the following subject areas in the 1992 NAEP assessment.

- Reading: Three newly developed IRT scales for the main assessment of reading and one IRT scale linking 1992 results to results from reading assessments in 1971, 1975, 1979, 1984, 1988, and 1990.
- Mathematics: Six scales linking back to the 1990 main mathematics assessment and a unidimensional IRT mathematics scale linking 1992 results to results from mathematics assessments in 1973, 1976, 1982, 1986, and 1990.
- Science: One unidimensional scale linking 1992 results to results from science assessments in 1969, 1973, 1977, 1982, 1986, and 1990.
- Writing: One newly developed polytomous IRT scale for the main assessment of writing and one newly developed polytomous scale linking writing results from the 1984, 1988, 1990, and 1992 assessments.

Details follow in Chapters 12, 13, 14, and 15.

Chapter 12

DATA ANALYSIS FOR THE READING ASSESSMENT[1]

John R. Donoghue, Steven P. Isham, Drew W. Bowker, and David S. Freund

Educational Testing Service

This chapter describes the analyses performed on the responses to the cognitive and background items in the 1992 assessment of reading. These analyses led to the results presented in *Trends in Academic Progress: Achievement of U.S. Students in Science, 1969 to 1992; Mathematics, 1973 to 1992; Reading, 1971 to 1992; and Writing, 1984 to 1992* (Mullis, Dossey, Campbell, Gentile, O'Sullivan, & Latham, 1994) and in the *NAEP 1992 Reading Report Card for the Nation and the States: Data from the National and Trial State Assessments* (Mullis, Campbell, & Farstrup, 1993). The emphasis of this chapter is on the methods and results of procedures used to develop the IRT-based scale scores that formed the basis of these reports. However, some attention is given to the analysis of constructed-response items as reported in the *NAEP 1992 Reading Report Card for the Nation and the States*. The theoretic underpinnings of the IRT and plausible values methodology described in this chapter are given in Chapter 11, and several of the statistics are described in Chapter 9.

The objectives of the reading analyses were to:

- prepare scale values and perform all analyses necessary to produce a long-term trend report in reading. The reading trend results include the years 1971, 1975, 1980, 1984, 1988, 1990, and 1992.

- prepare scale values for the analysis of the main focused-BIB reading samples.

- prepare the analysis of the oral reading assessment. The oral reading sample 9[Rdg-MainO] is a subset of the 9[Rdg-MainP] sample. Through merging of records, performance on the oral reading assessment was linked to the main reading scales. Analyses of the oral reading assessment are described in *Between the Lines: Listening to Children Read Aloud*, and analyses of the associated interview data are described in *Between the Lines: Interviewing Children About Their Literary Experiences* (Campbell, Kapinus, & Beatty, 1994). Therefore, the 9[Rdg-MainO] sample will not be discussed further in this chapter.

---

[1] Data analysis and scaling were performed by Drew Bowker, Steve Isham, and David Freund. Nancy Allen and Robert Mislevy consulted on IRT scaling and generation of plausible values.

The major analysis components are discussed in turn. Some aspects of the analysis, such as procedures for item analysis, scoring of constructed-response items, and methods of scaling, are described in previous chapters and are therefore not detailed here.

The student samples that were administered reading items in the 1992 assessment are shown in Table 12-1. (See Chapters 1 and 3 for descriptions of the target populations and the sample design used for the assessment.) Data from the first three samples (denoted Rdg-MainP) were used in the main analysis. The final three samples (denoted RW-LTTrend) were collected for long-term trend analyses.

Data from the 1992 main focused-BIB samples were scaled separately from the data from the long-term trend samples that contributed to the trends in reading achievement. Accordingly, the trend and main analyses are presented in separate sections. Section 12.1 pertains to the scaling of the data from the trend long-term trends; section 12.2 contains information about the scaling of the data from the main focused-BIB samples.

Table 12-1
NAEP 1992 Reading Student Samples

Sample	Booklets	Mode	Cohort Assessed	Time of Testing	Age Definition	Modal Grade	Number Assessed
9 [Rdg-MainP]	30-45	Print	Age 9/grade 4	Winter	CY	4	8,416
13 [Rdg-MainP]	30-49	Print	Age 13/grade 8	Winter	CY	8	14,942
17 [Rdg-MainP]	30-50	Print	Age 17/grade 12	Winter	CY	12	15,315
9 [Rdg-MainO]	N/A	------	Grade 4	Winter	CY	4	1,583
9 [RW-LTTrend]	51-56	Print	Age 9/grade 4	Winter	CY	4	7,062
13 [RW-LTTrend]	51-56	Print	Age 13/grade 8	Fall	CY	8	5,514
17 [RW-LTTrend]	51-56	Print	Age 17/grade 11	Spring	Not CY	11	5,569

LEGEND:

Rdg	Reading
RW	Reading and writing
MainP	Main assessment, print administration
MainO	Special oral reading assessment. Students participating in the oral reading assessment were a subsample of those in the 9[Rdg-MainP] print assessment
LTTrend	Long-term trend assessment
Print	Printed administration
CY	Calendar year: birthdates in 1982, 1978, and 1974 for ages 9, 13, and 17
Not CY	Age 17 only: birthdates between Oct. 1, 1974 and Sept. 30, 1975

12.1 LONG-TERM TREND DATA ANALYSIS

The trend results reported in *Trends in Academic Progress* are based on print administrations and occur at all of the age levels. The samples involved in the analysis are shown in Table 12-1 as 9[RW-LTTrend], 13[RW-LTTrend], and 17[RW-LTTrend]. The long-term trend booklets for these samples contained blocks of reading and writing items. All items were administered in print form. All students received a block of common background questions, distinct for each age, in addition to subject-area background questions, which were presented in the cognitive blocks. The booklets are identical to those used for long-term trend assessments in 1984, 1988, and 1990. The booklets and the blocks within those booklets are listed in Tables 4-1, 4-2, and 4-3 of Chapter 4. Additional information about all of the items in these blocks is in Tables 4-4, 4-5, and 4-6 of that chapter. This chapter includes specific information about the trend items that were scaled. Both age- and grade-selected students contributed to the trend scaling. However, only students in the "age-only" portion of the reading trend samples contributed to the results presented in *Trends in Academic Progress*.

Table 12-2 clarifies the relationships between the 1992 long-term trend samples and samples from previous years. For ages 9, 13 and 17, the [RW-LTTrend] samples allow direct comparisons with 1990 and 1988 samples, as well as with 1984 samples. The long-term trend scale was established in 1984. The 1971, 1975, and 1980 assessments were linked to the 1984 assessment using a complex equating strategy, which is described in *Implementing the New Design: The NAEP 1983-84 Technical Report* (Beaton, 1987). At each age, several intact booklets were retained from the 1984 assessment. These intact booklets form the basis of the reading trend assessment in 1988, 1990, and 1992. Information about the 1988 assessment is available in *Focusing the New Design: The NAEP 1988 Technical Report* (Johnson & Zwick, 1990), and information about the 1990 assessment is given in *The NAEP 1990 Technical Report* (Johnson & Allen, 1992).

The 1992 long-term trend included, at each age level, six of the assessment booklets administered in 1984. These booklets (51-56) contained both reading and writing blocks, as well as background items. Although these long-term trend booklets represented only about a tenth of the reading booklets administered in the complex 1984 BIB design,[2] they contained 10 of the 12 reading blocks that were scaled at each age/grade level in 1984. The samples of students who received these long-term trend booklets are described in Table 12-1 and in Chapter 4. The purpose of the long-term reading trend analysis was to add to the reading trend results that extended from 1971 to 1990 for ages 9, 13, and 17. The numbers of scaled items for each age are presented in Table 12-3. Each age was scaled separately. The numbers of items scaled in 1992 that were common across assessment years are given in Table 12-4. As was the case for previous trend analyses, the trend scale is univariate. Dimensionality analyses conducted following the 1984 assessment showed that the reading items were well summarized by a unidimensional scale (Zwick, 1987a).

The steps in the reading trend analysis are documented in the following sections. As is usual in NAEP analyses, the first step was to gather item and block information. The trend items were then calibrated according to the IRT model. Plausible values were generated after conditioning on available background variables. Finally, the scale values were placed on the final reading trend proficiency scale used in previous trend assessments.

---

[2] The long-term trend assessment included 1984 booklets 16, 17, 27, 34, 55, and 60 at age 9 and booklets 13, 16, 17, 21, 34, and 57 at ages 13 and 17 (see J. R. Johnson, 1987, pp. 120-121). The 1984 main assessment focused BIB design included 57 booklets that contained at least one scaled reading block at age 9 and 56 such booklets at ages 13 and 17.

Table 12-2
NAEP Reading Samples Contributing to 1992 Long-term Trend Results, 1971-1992

Cohort	Year	Sample	Subjects	Time of Testing	Mode of Administration	Age Definition	Modal Grade
Age 9	1971	Main	RL	Winter	Tape	CY	4
	1975	Main	RA	Winter	Tape	CY	4
	1980	Main	RA	Winter	Tape	CY	4
	1984	Main	RW	Winter, spring	Print	CY	4
	1984	T-84	RW	Winter	Tape	CY	4
	1988	LTTrend	RW	Winter	Print	CY	4
	1990	LTTrend	RW	Winter	Print	CY	4
	1992	LTTrend	RW	Winter	Print	CY	4
Age 13	1971	Main	RL	Fall	Tape	CY	8
	1975	Main	RA	Fall	Tape	CY	8
	1980	Main	RA	Fall	Tape	CY	8
	1984	Main	RW	Winter, spring	Print	CY	8
	1984	T-84	RW	Fall	Tape	CY	8
	1988	LTTrend	RW	Fall	Print	CY	8
	1990	LTTrend	RW	Fall	Print	CY	8
	1992	LTTrend	RW	Fall	Print	CY	8
Age 17	1971	Main	RL	Spring	Tape	Not CY	11
	1975	Main	RABS	Spring	Tape	Not CY	11
	1980	Main	RA	Spring	Tape	Not CY	11
	1984	Main	RW	Winter, spring	Print	Not CY	11
	1984	T-84	RW	Spring	Tape	Not CY	11
	1988	LTTrend	RW	Spring	Print	Not CY	11
	1990	LTTrend	RW	Spring	Print	Not CY	11
	1992	LTTrend	RW	Spring	Print	Not CY	11

LEGEND:

RL	Reading and literature
RA	Reading and art
RABS	Reading, art, index of basic skills
RW	Reading and writing
Main	Main assessment
T-84	Special sample in the 1984 assessment that was used to establish links to previous assessments (1971-1980) for the purposes of long-term trend
LTTrend	Long-term trend (these samples received common booklets within an age group)
Print	Print administration
Tape	Audiotape administration
CY	Calendar year: birthdates (1992 sample) in 1982 and 1978 for ages 9 and 13
Not CY	Age 17 only (1992 sample): birthdates between Oct 1, 1974 and Sept. 30, 1975

BEST COPY AVAILABLE

Table 12-3

Numbers of Scaled Reading Long-term Trend Items Common Across Ages

Age	Number of Items
9 only	60
13 only	21
17 only	23
9 and 13 only	13
9 and 17 only	2
13 and 17 only	42
9, 13, and 17	27
Total	188

Table 12-4

Numbers of Scaled Reading Long-term Trend Items Common Across Assessments

Assessment Year	Number of Items		
	Age 9	Age 13	Age 17
1984, 1990, 1992	101	101	93
1984, 1992	102	103	94
1984, 1988, 1990, 1992	98	98	87
1980, 1984, 1988, 1990, 1992	67	71	52
1975, 1980, 1984, 1988, 1990, 1992	36	45	37
1971, 1975, 1980, 1984, 1988, 1990, 1992	36	45	37

12.1.1 Item Analysis for the Long-term Trend Assessment

Conventional item analyses did not identify any difficulties with the long-term trend data. Table 12-5 contains the number of items, size of the sample administered the block, mean weighted proportion correct, mean weighted r-biserial, and mean weighted alpha as a measure of reliability for each block. Because the blocks were presented in self-paced, print-administered form, the weighted proportion of students attempting the last item is included in the table to give an indication of the speededness of each block. Common labeling of these blocks across ages does not denote common items. Student weights were used for all statistics, except for the sample sizes. The average values reflect only the items in the block that were scaled. The average weighted proportion correct tended to be somewhat higher than it was for 1990. Overall, however, the 1992 item-level statistics were not very different from those for the 1984, 1988, and 1990 assessments.

12.1.2 Treatment of Constructed-response Items

Data for constructed-response items in the trend analysis were used for the 1984, 1990, and 1992 assessments only. Constructed-response items were not included in the original scoring of the 1988 reading assessment because a previous study (Zwick, 1988) had shown that scoring inconsistencies (drops in interrater reliability and/or scorer drift—that is, scorers showed evidence of rating items more strictly or more leniently than did the original, 1984 scorers) had affected these items. A similar review was performed on the 1992 constructed-response items. In general, the 1992 scoring did not suffer from the same inconsistencies as the 1988 scoring. Therefore, most of the 1992 constructed-response items were used in the trend analysis.

At each age, several constructed-response items were found to exhibit drops in interrater reliability and/or scorer drift. Items exhibiting marked item drift or drops in interrater reliability were excluded from calibration. Three of the 198 total trend reading items were excluded. These items are listed in Table 12-6. The remaining constructed-response items were dichotomized according to criteria developed by subject-area experts. The dichotomized versions of the constructed-response items were included in the calibration. Four additional constructed-response items exhibited differences in interrater reliability from 1990 to 1992. To further ensure stability of the common calibration, these items were scaled separately for each administration (see section 12.1.3 below).

12.1.3 IRT Scaling

12.1.3.1 Item Parameter Estimation

The first step in the scaling process was the estimation of item parameters for the trend items. This item calibration was performed using the BILOG/PARSCALE program described in Chapter 11. Items were calibrated separately for each of the three age/grade groups. Item parameters were estimated using combined data from the assessment years 1990 and 1992, treating each assessment as a sample from a separate subpopulation. Student weights were used

Table 12-5

Descriptive Statistics for Item Blocks
Reading Long-term Trend Samples

Statistics	Blocks										
	BH	BJ	BK	BL	BM	BN	BO	BP	BQ	BR	BV
Age 9											
Number of scaled items	10	8	11	7	11	12	11	—	11	12	9
Number of scaled constructed-response items	1	0	0	1	1	1	0	—	0	0	3
Unweighted sample size	811	842	837	821	817	810	799	—	824	1606	814
Average weighted proportion correct	.59	.52	.42	.51	.43	.58	.49	—	.55	.47	.65
Average weighted r-biserial	.76	.70	.65	.83	.69	.77	.60	—	.73	.67	.76
Weighted alpha reliability	.77	.66	.74	.75	.75	.84	.63	—	.80	.77	.76
Weighted proportion of students attempting last item	.90	.92	.81	.76	.70	.71	.90	—	.86	.87	.97
Age 13											
Number of scaled items	12	9	8	5	11	12	10	9	16	11	—
Number of scaled constructed-response items	1	0	0	0	1	1	1	1	0	0	—
Unweighted sample size	678	671	658	649	657	640	672	662	648	676	—
Average weighted proportion correct	.64	.64	.66	.73	.60	.67	.65	.72	.63	.70	—
Average weighted r-biserial	.68	.65	.74	.89	.68	.71	.62	.77	.58	.72	—
Weighted alpha reliability	.64	.60	.68	.58	.68	.79	.53	.69	.72	.73	—
Weighted proportion of students attempting last item	.96	.94	.99	.98	.91	.76	.83	.92	.79	.98	—
Age 17											
Number of scaled items	12	5	8	6	11	12	13	10	10	7	—
Number of scaled constructed-response items	1	1	0	1	1	1	1	1	0	0	—
Unweighted sample size	714	730	737	751	737	760	746	730	749	714	—
Average weighted proportion correct	.71	.82	.80	.74	.69	.84	.67	.75	.58	.69	—
Average weighted r-biserial	.76	.94	.80	.84	.74	.80	.60	.75	.62	.83	—
Weighted alpha reliability	.74	.54	.64	.51	.66	.78	.70	.76	.64	.73	—
Weighted proportion of students attempting last item	.97	.96	1.00	.96	.98	.87	.69	.86	.93	1.00	—

BEST COPY AVAILABLE

Table 12-6

Items Deleted from the Reading Long-term Trend Analysis

Age	Block	Item	Reason for Exclusion
9	BJ	N001801	Never scaled, extremely low probability of correct response
	BM	N003003	Exclude, poor fit to IRT model
	BJ	N008905	Exclude, marked score drift
13	BJ	N001801	Exclude, poor fit to IRT model
	BJ	N001904	Exclude, marked score drift
	BK	N002302	Never scaled, nonordinal item
	BL	N002804	Exclude, low interrater reliability
	BQ	N005001	Exclude, poor fit to IRT model
17	BK	N002302	Never scaled, nonordinal item
	BQ	N015905	Exclude, marked score drift

for the calibration. To ensure that each assessment year had a similar influence on the calibration, student weights for 1990 examinees were multiplied by a constant, to adjust them to have the same sum as the sum of the weights for the 1992 examinees. Approximately 300-500 examinees were present in each assessment year for each item.

Starting values were computed from item statistics based on the 1992 data set. As described in Chapter 9, BILOG/PARSCALE calibrations were done in two stages. At stage one, the proficiency distribution of each assessment year was constrained to be normally distributed, although the means and variances differed across assessment year. The values of the item parameters from this normal solution were then used as starting values for a second stage estimation run in which the proficiency distribution (modeled as a separate multinomial distribution for each assessment year) was estimated concurrently with item parameters. Calibration was concluded when changes in item parameters became negligibly small (i.e., less than .005).

12.1.3.2 Evaluation of Model Fit

During and subsequent to item parameter estimation, evaluations of the fit of the IRT models were carried out for each of the items. These evaluations were based primarily on graphical analysis. First, model fit was evaluated by examining plots of nonmodel-based estimates of the expected proportion correct (conditional on proficiency) versus the proportion correct predicted by the estimated item response function (see Chapter 9 and Mislevy and Sheehan, 1987, p. 302). In making decisions about excluding items from the final scales, a balance was sought between being too stringent, hence deleting too many items and possibly damaging the content representativeness of the pool of scaled items, and being too lenient, hence including items with model fit poor enough to endanger the types of model-based inferences made from NAEP results. For the majority of the items, the model fit was extremely good. Items that clearly did not fit the model were not included in the final scales; however, a certain degree of misfit was tolerated for a number of items included in the final scales (see section 12.2.2 for example item plots).

In addition to the constructed-response items dropped due to poor interrater reliability and item drift, one item was deleted from the scaling of the age 9 trend data, and two items were deleted from the age 13 trend analysis, due to poor fit to the IRT model. The items were too difficult for these students to yield reliable estimates of item parameters. No other trend items were deleted from the 1992 trend analysis. Table 12-6 lists items that were excluded from the long-term trend assessment. These items will be re-examined in future trend assessments.

The adequacy of the assumption of a common item response function across assessment years was also evaluated. This was evaluated by comparing the nonmodel-based expected proportions for each assessment year to the single, model-based item response function fit by BILOG/PARSCALE. Items that showed clear evidence of functioning differently across assessments were treated as separate items for each assessment year—that is, separate item response functions were estimated for each assessment. As was the case with deleting items, in making decisions about scaling items separately by assessment year, a balance was sought between being too stringent, hence splitting too many items and possibly damaging the common item link between the assessment years, and being too lenient, hence including items with model fit poor enough to endanger the model-based trend inferences. These items will be re-examined in future trend assessments.

At each age, several items were calibrated separately for each assessment year, for one of two reasons. First, as noted above, some constructed-response items were calibrated separately across assessments due to changes in rater reliability and/or rater drift across assessment years. Second, examination of residual plots identified some items as functioning differently across assessments. Figure 12-1 shows item N003704 from the analysis for age 17/grade 12. Data are presented for 1990 (rectangles) and for 1992 (ovals). For proficiency values less than 0.5, the two sets of symbols diverge, and the discrepancy is substantial for $\theta < 0.0$. The top (1990 data) and bottom (1992 data) of Figure 12-2 show the plots for the item treated separately by assessment year. The remaining misfit is relatively small, and affects a small proportion of the population. Overall, 7 of the 195 long-term trend reading items were calibrated separately by assessment year. Table 12-7 lists the items that were calibrated separately across assessment years.

Table 12-7
Items Calibrated Separately by Assessment Year in the Reading Long-term Trend Analysis

Age	Block	Item	Reason for Separate Calibration
9	BH	N001507	Constructed-response item, change in interrater reliability
	BN	N003704	Poor fit across assessments to common item response function
13	BH	N001507	Constructed-response item, change in interrater reliability
	BM	N003104	Constructed-response item, change in interrater reliability
17	BH	N001507	Poor fit across assessments to common item response function
	BL	N002804	Constructed-response item, change in interrater reliability
	BN	N003704	Poor fit across assessments to common item response function

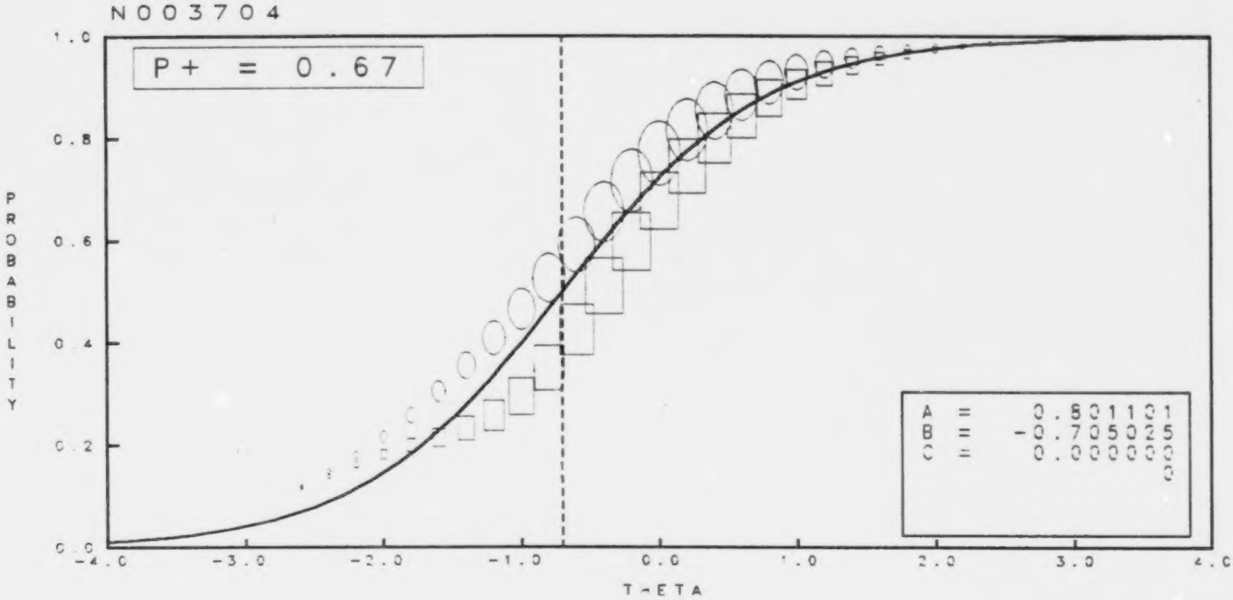
A list of the items scaled for each of the ages, along with their item parameter estimates, appears in Appendix E.

12.1.4 Generation of Plausible Values

The generation of plausible values was conducted independently for each age/grade level for each of the assessment years. The item parameters from BILOG/PARSCALE, final student weights, item responses, and selected background variables were used with the computer program BGROUP (described in Chapter 11) to generate the values for each age. The background variables included student demographic characteristics (e.g., race/ethnicity of the student, highest level of education attained by parents), students' perceptions about reading, and student behavior both in and out of school (e.g., amount of television watched daily, amount of homework done each day). Appendix F gives the codings for the conditioning variables and the estimated conditioning effects for the three age groups. The estimated conditioning effects in

Figure 12-1

Example Long-term Trend Item (N003704, Age 17)
Demonstrating Differential Item Functioning Across Assessment Years 1990 and 1992*

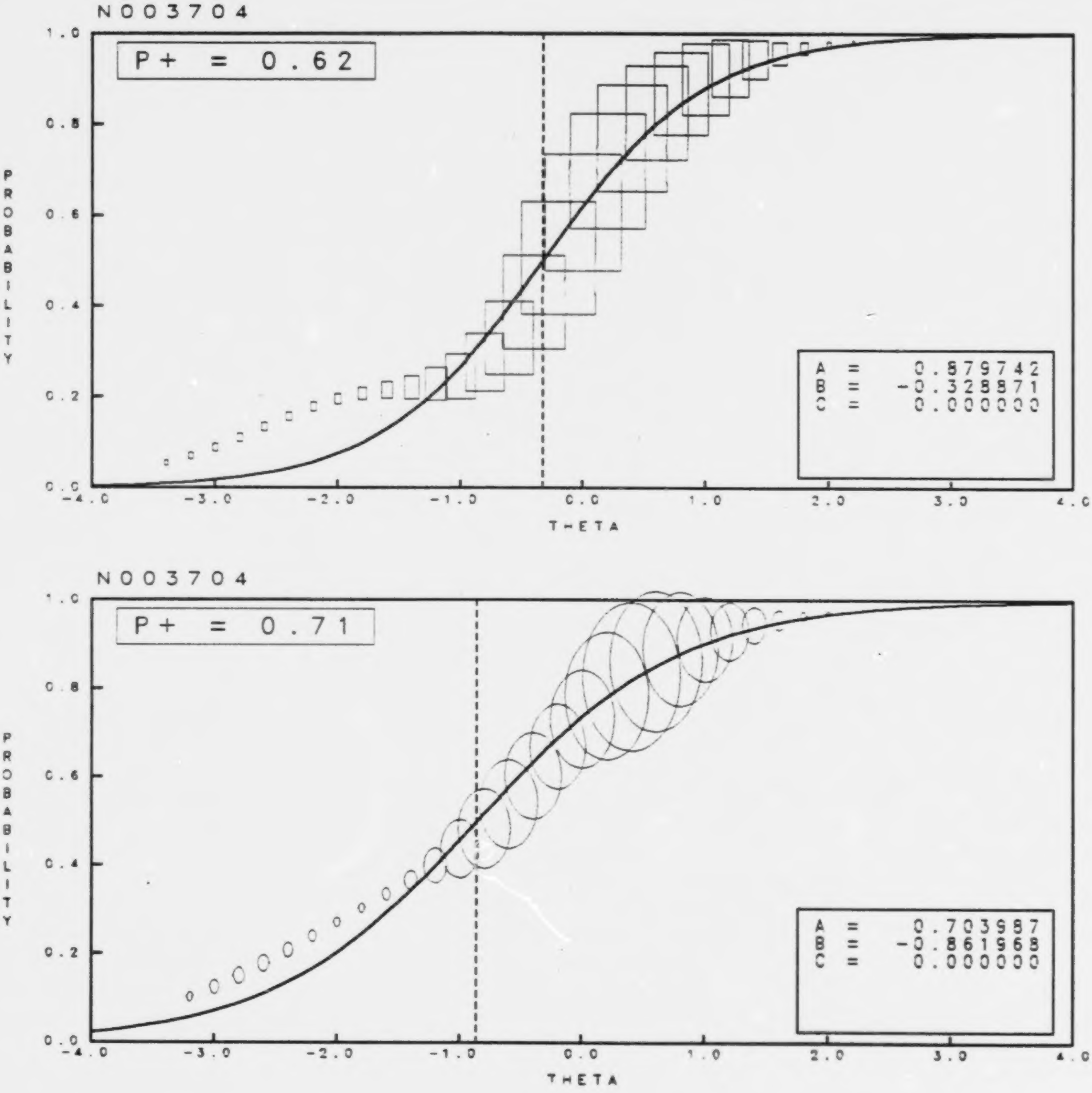


---

* Solid line represents common, model-based trace; rectangles represent 1990 data; ovals represent 1992 data.

Figure 12-2

Example Long-term Trend Item (N003704, Age 17)
Fitting Separate Item Response Functions for Each Assessment Year*



---

* Rectangles represent 1990 data; ovals represent 1992 data.

the tables are expressed on the scale of the original BILOG/PARSCALE calibration. Table 12-8 contains a list of the number of background variables included in conditioning, as well as the proportion of variance accounted for by the conditioning model for each age/grade.

Table 12-8
Proportion of Proficiency Variance Accounted for by the Conditioning Model
for the Reading Long-term Trend Assessment

Age/Grade	Number of Conditioning Variables	Proportion of Proficiency Variance
9/4	41	.372
13/8	41	.368
17/11	39	.399

12.1.5 The Final Long-term Trend Proficiency Scale

The linear indeterminacy of the long-term trend scale was resolved by linking the 1992 trend scales to previous trend scales. For each age, the item parameters from the joint calibration based on data from 1990 and 1992 were used with the 1990 data to find plausible values for the 1990 data. The mean and standard deviation of all of the plausible values were calculated and matched to the mean and standard deviation of all of the plausible values based on the original analysis of the 1990 data, as given in earlier reports. The transformations that resulted from this matching of the first two moments for the 1990 data are

Age 9: $\theta_{\text{proficiency}} = 40.86 \bullet \theta_{\text{calibrated}} + 210.54$

Age 13: $\theta_{\text{proficiency}} = 38.75 \bullet \theta_{\text{calibrated}} + 256.93$

Age 17: $\theta_{\text{proficiency}} = 42.19 \bullet \theta_{\text{calibrated}} + 287.73$

where $\theta_{\text{proficiency}}$ denotes values on the final transformed scale and $\theta_{\text{calibrated}}$ denotes values on the calibration scale. Overall summary statistics for the trend samples are given in Table 12-9.

As in the past, interpretation of the trend results was facilitated through the provision of scale anchoring information. In 1984, five NAEP reading scale levels were selected as anchor points. These points (described in *Trends in Academic Progress*) are:

150 - Simple, discrete reading tasks;
200 - Partially developed skills and understanding;
250 - Interrelation of ideas and generalizations;
300 - Understanding complicated information; and
350 - Learning from specialized reading materials.

Table 12-9

Means and Standard Deviations on the Reading Long-term Trend Proficiency Scale

		All Five Plausible Values	
Age	**Assessment**	**Mean**	**S. D.**
9	1984	211.0	41.1
	1988	211.8	41.2
	1990	209.2	44.7
	1992	210.5	40.4
13	1984	257.1	35.5
	1988	257.5	34.7
	1990	256.8	36.0
	1992	259.8	39.4
17	1984	288.8	40.3
	1988	290.1	37.1
	1990	290.2	41.3
	1992	289.7	43.0

Detailed descriptions of the skills required to read at each level were derived and benchmark exercises were selected to exemplify each level. These same anchor points were used in the 1988, 1990, and 1992 reading trend reports. The estimated proportion of students in each reporting category who are at or above each anchor point were examined in *Trends in Academic Progress.*

12.1.6 Partitioning of the Estimation Error Variance

For each age, the variance of the final, transformed scale mean was partitioned as described in Chapter 11. This analysis yielded estimates of the proportion of error variance due to sampling students and the proportion due to the latent nature of θ. These estimates are given in Table 12-10 (for stability of the estimates, they are based on 100 plausible values). More detailed information is available for gender and race/ethnicity subgroups in Appendix N.

Table 12-10
Estimation Error Variance and Related Coefficients for the Reading Long-term Trend Assessment

Age	Total Estimation Error Variance	Proportion of Variance Due to...	
		Student Sampling	Latency of θ
9	0.73	0.88	0.12
13	1.30	0.91	0.09
17	1.20	0.86	0.14

12.2 MAIN DATA ANALYSIS

12.2.1 Overview

This section describes the analyses carried out in the development of the 1992 main reading scales. The procedures used were similar to those employed in the analysis of the 1990 assessment of mathematics (Yamamoto & Jenkins, 1992) and parallel those used in the analysis of the 1992 Trial State Assessment in reading (Allen, Mazzeo, Isham, Fong, & Bowker, 1993), and are based on the philosophical and theoretical underpinnings described in Chapter 11.

There were six major steps in the analysis of the 1992 main reading assessment, each of which is described in a separate section:

- conventional item and test analysis (section 12.2.3.1);
- differential item functioning analysis (section 12.2.3.2);
- item response theory (IRT) scaling (section 12.2.4);

- estimation of national and subgroup proficiency distributions based on the "plausible values" methodology (section 12.2.5);

- transformation of the purposes-of-reading scales to the proficiency metric (section 12.2.6.1); and

- creation of the composite reading scale (section 12.2.6.2).

12.2.2 Description of Items and Assessment Booklets

The 1992 NAEP main assessment differed from the long-term trend assessment in regard to the sample age definition, time of testing, the objectives that define the emphasis of the assessment, and most of the items used. It also differed from the 1988 and 1990 main NAEP assessments in the same regards. Because of these differences, equating the main and the long-term trend assessments with the method applied in the previous section is not appropriate. Neither is a direct comparison to the 1990 main assessment. The 1992 main reading assessment can be used to start a new baseline for measuring trends in the nation.

The items in the assessment were based on the curriculum framework described in *Reading Framework for the 1992 National Assessment of Educational Progress* (National Assessment Governing Board, 1992b). Compared to previous NAEP assessments, the 1992 assessment contains longer reading passages that are intended to be more authentic examples of the reading tasks encountered in and out of school. As described in the reading framework, these blocks are organized into three scales, corresponding to three purposes of reading: Reading for Literary Experience, Reading to Gain Information, and Reading to Perform a Task. At grade 4/age 9, only the first two purposes are represented. Scales were produced for each of the purposes of reading. In addition, a composite scale for reading was created as a weighted sum of the purposes-of-reading scales.

The data from the main focused-BIB assessment of reading (9[Rdg-MainP], 13[Rdg-MainP], and 17[Rdg-MainP]) were used for main analyses comparing the levels of reading achievement for various subgroups of the 1992 target populations. The main assessment included three student cohorts: students who were either in the fourth grade or 9 years old, students who were either in the eighth grade or 13 years old, and students who were either in the twelfth grade or 17 years old. The birth date ranges for age-eligible students were based on the 1982, 1978, and 1974 calendar years respectively for ages 9, 13, and 17. The sampled students in each of these three cohorts were assessed in the winter. The samples in the main assessment are listed in Table 12-1.

In the main samples, each student was administered a booklet containing two blocks of cognitive reading items, a block of background questions common to all booklets for a particular age/grade level, a block of questions concerning the student's motivation and his or her perception of the difficulty of the cognitive items, and a block of reading-related background questions common to all reading booklets for a particular age/grade level. Eight (grade 4) or nine (grades 8 and 12) 25-minute blocks of reading items were administered at each age/grade level. As described in Chapter 2, the 25-minute blocks were combined into booklets according to a partially balanced incomplete block design. (See Chapter 4 for more information about the blocks and booklets.) In addition, 50-minute reading blocks were presented to the older students, one at grade 8 and two at grade 12. The 50-minute blocks were closely examined to insure the appropriateness of including them with the shorter blocks in the scaling.[3] Both age- and grade-selected students contributed to the main scaling. However, the "grade-only" portion of the main focused-BIB reading samples contributed to the means and percentages of the main results that are reported in the *NAEP 1992 Reading Report Card.*

Each block consisted of one or two reading passages, followed by several items. In addition to multiple-choice items, each block contained a number of constructed-response items, accounting for well over half of the testing time. Constructed-response items were scored by specially trained readers (described in Chapter 7). Some of the constructed-response items required only a few sentences or a paragraph response. These short constructed-response items were scored dichotomously as correct or incorrect. In addition, each block contained at least one extended constructed-response item, which required a more in-depth, elaborated response. These extended constructed-response items were scored polytomously on a five point (0-4) scale. During the scaling process, the 0 category (omitted and off-task responses) was recoded. Off-task responses were treated as "not administered" for each of the items, and omitted responses were combined with the next lowest category, "Unsatisfactory". This resulted in a 0-to-3 scale:

0 - Unsatisfactory (and omit);
1 - Partial;
2 - Essential;
3 - Extensive, which demonstrates more in-depth understanding.

In addition, categories of a small number of items were combined ("collapsed"). These changes were made so that the scaling model used for these items fit the data more closely, and are described more fully in section 12.2.4.

The composition of each block of items, in terms of format and content, is given in Table 12-11. Common labeling of these blocks across age/grade levels does not denote common items. The numbers of items scaled in 1992 for each age/grade are presented in Table 12-12.

12.2.3 Item Analyses

12.2.3.1 Conventional Items and Test Analyses

Tables 12-13, 12-14, and 12-15 show the number of items in the block, the average weighted item score, average weighted polyserial correlation, and the weighted alpha reliability for each block administered. These statistics are described in Chapter 9. These values were calculated for the items within each block that was used in the scaling process. The tables also give the number of students who were administered the block and the percent not reaching the last item in the block. These numbers include only those students in the grade-only portion of the samples that contributed to the summary statistics provided in the *NAEP 1992 Reading*

---

[3] These analyses are described in *Assessing the Properties of Longer Blocks in the 1992 NAEP Reading Assessment* (Donoghue & Mazzeo, 1993).

Table 12-11
Item Type Composition of Blocks Used in Reading Main Assessment

Grade	Type of Item	Block										
		RC	RD	RE	RF	RG	RH	RI	RJ	RK	RM	RN
4	Number of multiple-choice items	6	5	7	5	4	5	4	6	—	—	—
	Number of short constructed-response items	4	6	3	4	5	4	4	5	—	—	—
	Number of extended constructed-response items	1	1	1	1	1	1	1	1	—	—	—
	Total	11	12	11	10	10	10	9	12	—	—	—
8	Number of multiple-choice items	4	8	7	5	6	7	7	5	3	5	—
	Number of short constructed-response items	6	4	3	5	6	6	4	6	8	6	—
	Number of extended constructed-response items	1	1	1	2	1	1	1	1*	1	2	—
	Total	11	13	11	12	13	14	12	12	12	13	—
12	Number of multiple-choice items	4	3	6	5	6	3	7	5	7	10	7
	Number of short constructed-response items	6	5	6	5	6	6	4	6	7	4	2
	Number of extended constructed-response items	1	1	0*	2	0*	1	1	1*	1	2	3
	Total	11	9	12	12	12	10	12	12	15	16	12

---

* This block contained one additional extended constructed-response item. The categories of that item were combined such that the version used in scaling had only two categories (see section 12.2.3). Therefore, the item is counted as "short constructed-response," reflecting the amount of information it contributed to the final scale.

Table 12-12

Numbers of Scaled Reading Items Common Across Grade Levels

Grade	**Reading for Literary Experience**	**Reading to Gain Information**	**Reading to Perform a Task**	**Total**
4 only	32	32	0	64
8 only	24	30	12	66
12 only	32	50	15	97
4 and 8 only	11	10	0	21
8 and 12 only	11	12	24	47
4 and 12 only	0	0	0	0
4, 8, and 12	0	0	0	0
Total	110	134	51	295

Table 12-13

Descriptive Statistics for Item Blocks by Position Within Booklet and Over All Occurrences
Reading Main Sample, Grade 4

Statistic	Block Position	Block							
		Reading for Literary Experience				Reading to Gain Information			
		RC	RD	RE	RI	RF	RG	RH	RJ
Number of scaled items	—	11	12	11	9	10	10	10	12
Unweighted sample size	1	782	773	781	761	784	787	793	798
	2	783	779	766	767	773	771	797	787
	ALL	1565	1552	1547	1528	1557	1558	1590	1585
Average weighted item score	1	.63	.68	.46	.48	.56	.44	.52	.64
	2	.63	.65	.44	.46	.55	.43	.49	.61
	ALL	.63	.67	.45	.47	.55	.43	.51	.63
Average weighted r-polyserial	1	.69	.69	.62	.70	.58	.69	.62	.63
	2	.75	.69	.63	.72	.63	.70	.62	.63
	ALL	.72	.69	.62	.71	.61	.69	.62	.63
Weighted alpha reliability	1	.80	.80	.74	.77	.67	.75	.70	.78
	2	.80	.79	.72	.76	.72	.76	.68	.74
	ALL	.80	.79	.73	.76	.69	.75	.69	.76
Weighted proportion of students attempting last item	1	.65	.58	.70	.68	.60	.55	.70	.75
	2	.85	.74	.82	.81	.82	.69	.77	.86
	ALL	.75	.66	.76	.75	.71	.62	.73	.81

Table 12-14

Descriptive Statistics for Item Blocks by Position Within Booklet and Over All Occurrences
Reading Main Sample, Grade 8

Statistic	Block Position	Block									
		Reading for Literary Experience			Reading to Gain Information				Reading to Perform a Task		
		RC	RD	RE	RF	RG	RH	RM*	RI	RJ	RK
Number of scaled items	—	11	13	11	12	13	14	13	12	12	12
Unweighted sample size	1	859	853	860	820	890	866	—	836	870	851
	2	834	857	829	876	864	846	—	849	862	848
	ALL	1693	1710	1689	1696	1754	1712	1731	1685	1732	1699
Average weighted item score	1	.40	.59	.69	.55	.68	.57	—	.48	.57	.64
	2	.39	.57	.66	.52	.66	.54	—	.47	.55	.62
	ALL	.39	.58	.68	.53	.67	.56	.50	.48	.56	.63
Average weighted r-polyserial	1	.67	.65	.74	.61	.72	.64	—	.62	.60	.68
	2	.66	.68	.74	.65	.74	.68	—	.61	.63	.71
	ALL	.66	.67	.74	.63	.73	.66	.59	.62	.61	.69
Weighted alpha reliability	1	.73	.76	.78	.71	.80	.78	—	.65	.69	.79
	2	.72	.78	.79	.75	.81	.80	—	.64	.73	.80
	ALL	.72	.77	.78	.73	.80	.79	.72	.65	.71	.80
Weighted proportion of students attempting last item	1	.74	.79	.93	.80	.80	.74	—	.95	.76	.75
	2	.81	.81	.94	.81	.86	.84	—	.94	.83	.87
	ALL	.78	.80	.94	.81	.83	.79	.75	.95	.79	.81

---

* A 50-minute block that comprised an entire booklet.

Table 12-15

Descriptive Statistics for Item Blocks by Position Within Booklet and Over All Occurrences
Reading Main Sample, Grade 12

Statistic	Block Position	Block										
		Reading for Literary Experience			Reading to Gain Information					Reading to Perform a Task		
		RC	RD	RE	RF	RG	RH	RM*	RN*	RI	RJ	RK
Number of scaled items	——	11	9	12	12	12	10	16	12	12	12	15
Unweighted sample size	1	764	749	746	752	752	723	—	—	744	781	741
	2	746	710	747	726	743	765	—	—	751	753	754
	ALL	1510	1459	1493	1478	1495	1488	1513	1533	1495	1534	1495
Average weighted item score	1	.54	.62	.59	.68	.67	.51	—	—	.63	.72	.60
	2	.52	.57	.57	.68	.64	.50	—	—	.61	.70	.61
	ALL	.53	.60	.58	.68	.66	.50	.67	.60	.62	.71	.61
Average weighted r-polyserial	1	.68	.65	.64	.63	.60	.62	—	—	.64	.59	.54
	2	.69	.68	.67	.66	.63	.65	—	—	.62	.62	.59
	ALL	.68	.66	.66	.65	.62	.63	.63	.59	.63	.61	.57
Weighted alpha reliability	1	.75	.68	.73	.67	.65	.65	—	—	.64	.65	.66
	2	.76	.68	.77	.67	.67	.70	—	—	.64	.71	.70
	ALL	.75	.68	.75	.67	.66	.68	.79	.71	.64	.68	.68
Weighted proportion of students attempting last item	1	.81	.53	.85	.87	.73	.87	—	—	.98	.81	.92
	2	.85	.66	.91	.91	.81	.91	—	—	.96	.83	.95
	ALL	.83	.59	.88	.89	.77	.89	.95	.97	.97	.82	.94

---

* A 50-minute block that comprised an entire booklet.

Report Card. Student weights were used for all statistics, except for the sample sizes. The results for the blocks administered to each age/grade level indicate that the blocks differ in number of items, average difficulty, reliability, and percent not reaching the last item, and so are not parallel to one another. Preliminary item analyses for all items within a block were completed before scaling; however, the results shown here indicate the characteristics of the items that contributed to the final scale.

As described in Chapter 9, in NAEP analyses (both conventional and IRT-based), a distinction is made between missing responses at the end of each block (not-reached) and missing responses prior to the last observed response (omitted). Items that were not reached were treated as if they had not been presented to the examinee, while omitted items were regarded as incorrect. The proportion of students attempting the last item of a block (or, equivalently, 1 minus the proportion not reaching the last item) is often used as an index of the degree of speededness of the block of items.

Standard practice at ETS is to treat all nonrespondents to the last item as if they had not reached the item. For multiple-choice and short constructed-response items, this convention produced a reasonable pattern of results, in that the proportion reaching the last item does not differ markedly from the proportion attempting the next to last item. However, for the blocks that ended with extended constructed-response items, this convention resulted in an implausibly large drop in the number of students attempting the final item. Therefore, for blocks that ended with an extended constructed-response item, students who attempted the next-to-last item but did not respond to the last item were classified as having intentionally omitted that item.

The results in Tables 12-13 to 12-15 indicate that the difficulty and internal consistency of the blocks varied. Such variability is expected, because the blocks were not constructed to be parallel. Based on the proportion of students attempting the last item, all of the blocks appear to be somewhat speeded. This effect is larger for grade 4 than for the other grades.

Small but consistent differences were noted based upon whether a block appeared first or second within a booklet. When the block appeared first in the booklet, the average item score tended to be higher and the average polyserial correlation tended to be lower. The largest differences were noted in the proportion of students not attempting the last item in the block; more students attempted the last item when the block appeared in the second position. It appears that students learned to pace themselves through the second block, based on their experience with the first block. A study was completed to examine the effect of the serial position differences. Due to the balance of the partial BIB design of the booklets, the serial position differences were found to have minimal effects on scaling.

12.2.3.2 Scoring the Constructed-response Items

As indicated earlier, the reading assessment included constructed-response items. Responses to these items were included in the scaling process. In addition, detailed analyses of the ordinal responses to the constructed-response items were also conducted, and are summarized in the *NAEP 1992 Reading Report Card*. Chapter 7 provides the ranges for percent agreement between raters for the items as they were originally scored. The percent agreement for the raters and Cohen's Kappa are given in Appendix K.

12.2.3.3 Differential Item Functioning

Prior to scaling, differential item functioning (DIF) analyses were conducted on the reading data for grades 4, 8, and 12, and for the Trial State Assessment sample at grade 4. The purpose of these analyses was to identify items which were differentially difficult for various subgroups and to reexamine such items, to determine their fairness and the appropriateness of including them in the scaling. The information in this section focuses mainly on the analyses conducted on the national data; results for the analyses of data from the Trial State Assessment are described in the technical report for that assessment.

DIF analyses were based upon the modification of the Mantel-Haenszel procedure as adapted by Holland and Thayer (1988), and described in Chapter 9. Currently, the Mantel-Haenszel procedure may only be used for dichotomous data. Thus, for the purposes of DIF analyses *only*, the extended constructed-response items were dichotomized—unsatisfactory and partial responses were treated as incorrect, and essential and extensive responses were treated as correct.

The "grade-only" portion of the main focused-BIB was used for DIF analyses. Sample sizes were large enough to compare male and female students, White and Black students, and White and Hispanic students. Weights were rescaled separately for each comparison, as described in Chapter 9. DIF analyses were conducted separately by grade. A given item was subjected to at least three, and as many as nine, separate DIF analyses. Table 12-16 summarizes the results of DIF analyses. The DIF index generated by the Mantel-Haenszel procedure is commonly used to ETS to place items into one of three categories: A, B, or C. "A" items exhibit no DIF, while "C" items exhibit a strong indication of DIF and should be examined more closely. Positive values of the index indicate items that are differentially easier for the "focal" group (female, Black, or Hispanic students) than for the "reference" groups (male or White students). Similarly, negative values indicate items that are differentially harder for the focal group than the reference group. An item that was classified as a "C" item in *any* analysis was considered to be a "C" item. Sixteen "C" items were identified in the main reading assessment.

Following standard practice at ETS for DIF analyses conducted on final forms, all "C" items were reviewed by a committee of trained test developers and subject-matter specialists. As described in Chapter 9, such committees are charged with making judgments about whether or not the differential difficulty of an item is *unfairly* related to group membership. The committee assembled to review NAEP items included both ETS staff and outside members with expertise in the field. The committee carefully examined each "C" item to determine if either the language or contents would tend to make the item more difficult for an identified group of examinees. It was the committee's judgment that none of the "C" items for the national or the Trial State Assessment data were functioning differentially due to factors irrelevant to test objectives. Hence, none of the items were removed from scaling due to differential item functioning.

Table 12-16

DIF Category by Grade

Grade	DIF Category	Analysis		
		Male/Female	**White/Black**	**White/Hispanic**
4	C-	0	1	0
	B-	3	1	1
	A-	34	48	37
	A+	43	31	43
	B+	5	3	4
	C+	0	1	0
8	C-	1	2	1
	B-	4	4	2
	A-	52	53	54
	A+	54	54	62
	B+	9	10	3
	C+	3	0	1
12	C-	2	0	0
	B-	12	12	3
	A-	49	47	53
	A+	52	62	69
	B+	16	11	7
	C+	2	1	1

---

* A = no indication of DIF; B = Weak indication of DIF; C = strong indication of DIF.

12.2.4 IRT Scaling

12.2.4.1 Item Parameter Estimation

Separate IRT-based scales were developed for each of the purposes of reading identified in the reading framework. As described in Chapter 11, multiple-choice items were fit using a 3PL model. Short constructed-response items were fit using a 2PL model. Extended constructed-response items were fit using the generalized partial credit model.

For calibration, all items that were not reached were treated as if they were not presented to the examinees. Recall that responses to extended constructed-response items that were off-task were also treated as if they had not been presented. The treatment of omitted responses differed according to the item type. Omitted responses to multiple-choice items were treated as fractionally correct. Omitted responses to short constructed-response items were treated as incorrect, and omitted responses to extended constructed-response items were assigned to the lowest category.

Item parameters were estimated using combined data from all three grade levels for the Reading to be Informed and Reading for Literary Experience scales; the Reading to Perform a Task scale was administered only at grades 8 and 12. Items that were administered at more than one grade (cross-grade items) were constrained to have equal item response functions across grades. However, a few items exhibited clear evidence of functioning differently across age/grade samples (see discussion in section 12.2.4.2). These items were treated as separate items for each age/grade level.

The calibration was performed using all the available examinees. Student sampling weights were used for the analysis. For scaling, sampling weights were restandardized to ensure that each age/grade sample had a similar sum of weights, and so had approximately equal influence in the calibration. Each grade was treated as a sample from a separate subpopulation. Thus separate proficiency distributions were estimated for each age/grade.

Item responses were calibrated using the BILOG/PARSCALE program. Starting values were computed from item statistics based on the entire data set. BILOG/PARSCALE calibrations were done in two stages. At stage one, the proficiency distribution of each age/grade was constrained to be normally distributed, although the means and variances differed across age/grades. The values of the item parameters from this normal solution were then used as starting values for a second stage estimation run in which the proficiency distribution (modeled as a separate multinomial distribution for each age/grade) was estimated concurrently with item parameters. Calibration was concluded when changes in item parameters became negligibly small.

12.2.4.2 Evaluation of Model Fit

During and subsequent to item parameter estimation, evaluations of the fit of the IRT models were carried out for each of the items. As with the long-term trend analysis, model fit of dichotomous items was evaluated by examining plots of nonmodel-based estimates of the expected proportion correct (conditional on proficiency) versus the proportion correct predicted by the estimated item response function. For extended constructed-response items, similar plots were produced for each item category response function (see Chapter 9).

For the majority of the items, the model fit was extremely good. In making decisions about excluding items from the final scales, a balance was sought between being too stringent, hence deleting too many items and possibly damaging the content representativeness of the pool of scaled items, and being too lenient, hence including items with model fit poor enough to endanger the types of model-based inferences made from NAEP results. Items that clearly did not fit the model were not included in the final scales; however, a certain degree of misfit was tolerated for a number of items included in the final scales.

A few polytomous items received special treatment in the scaling. Figure 12-3 shows one such item, R012111 from the Reading for Literary Experience scale at age 9/grade 4. There is a lack of fit for both the unsatisfactory and partial categories for low to moderate proficiency values. Categories 0 and 1 of this item were collapsed:

0 - Unsatisfactory
0 - Partial
1 - Essential
2 - Extensive

Figure 12-4 shows the recoded version of R012111 from the final scaling. The fit is substantially improved. Table 12-17 lists polytomous items that were recoded for scaling.

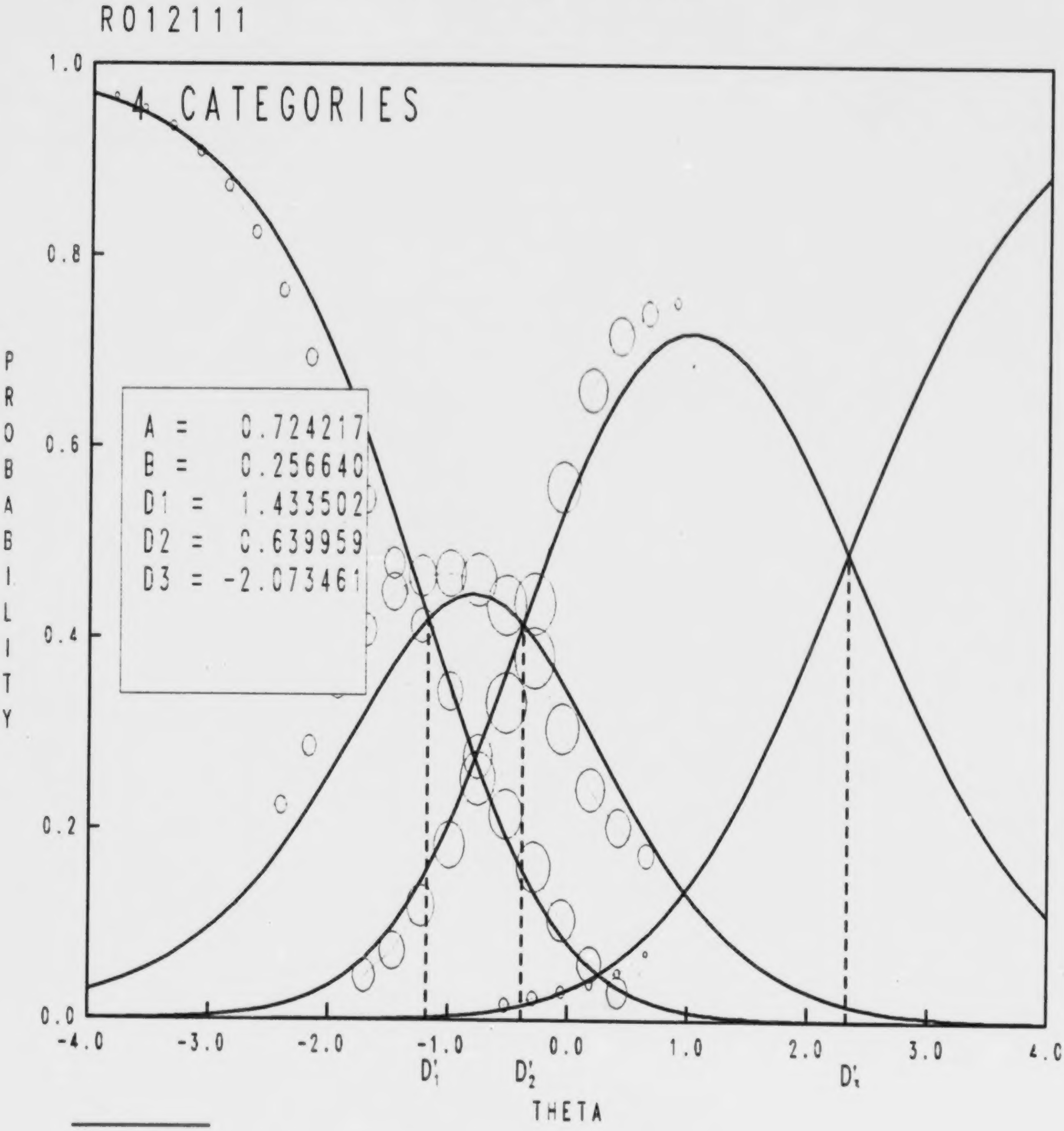
Table 12-17
Recoding of Polytomous Items for Scaling

Scale	NAEP ID	Block	Grade(s) Affected	Action
Reading for Literary Experience	R012111	RD	4	Combine 0 + 1 categories
	R013506	RD	12	Combine 0 + 1 categories
	R013610	RE	12	Combine 0 + 1, 2 + 3 categories (dichotomize)
Reading to Gain Information	R013706	RG	12	Combine 0 + 1, 2 + 3 categories (dichotomize)
	R013805	RH	12	Combine 0 + 1 categories
Reading to Perform a Task	R013004	RK	8	Combine 0 + 1 categories
	R013404	RJ	8, 12	Combine 0 + 1 categories
	R013406	RJ	8, 12	Combine 0 + 1, 2 + 3 categories (dichotomize)
	R013915	RK	12	Combine 0 + 1 categories

In addition, for cross-grade items, the adequacy of the assumption of a common item response function across grades was evaluated by comparing the nonmodel-based expected

Figure 12-3

Example Cross-sectional Polytomous Item (R012111, Age 9) Demonstrating Poor Model Fit*

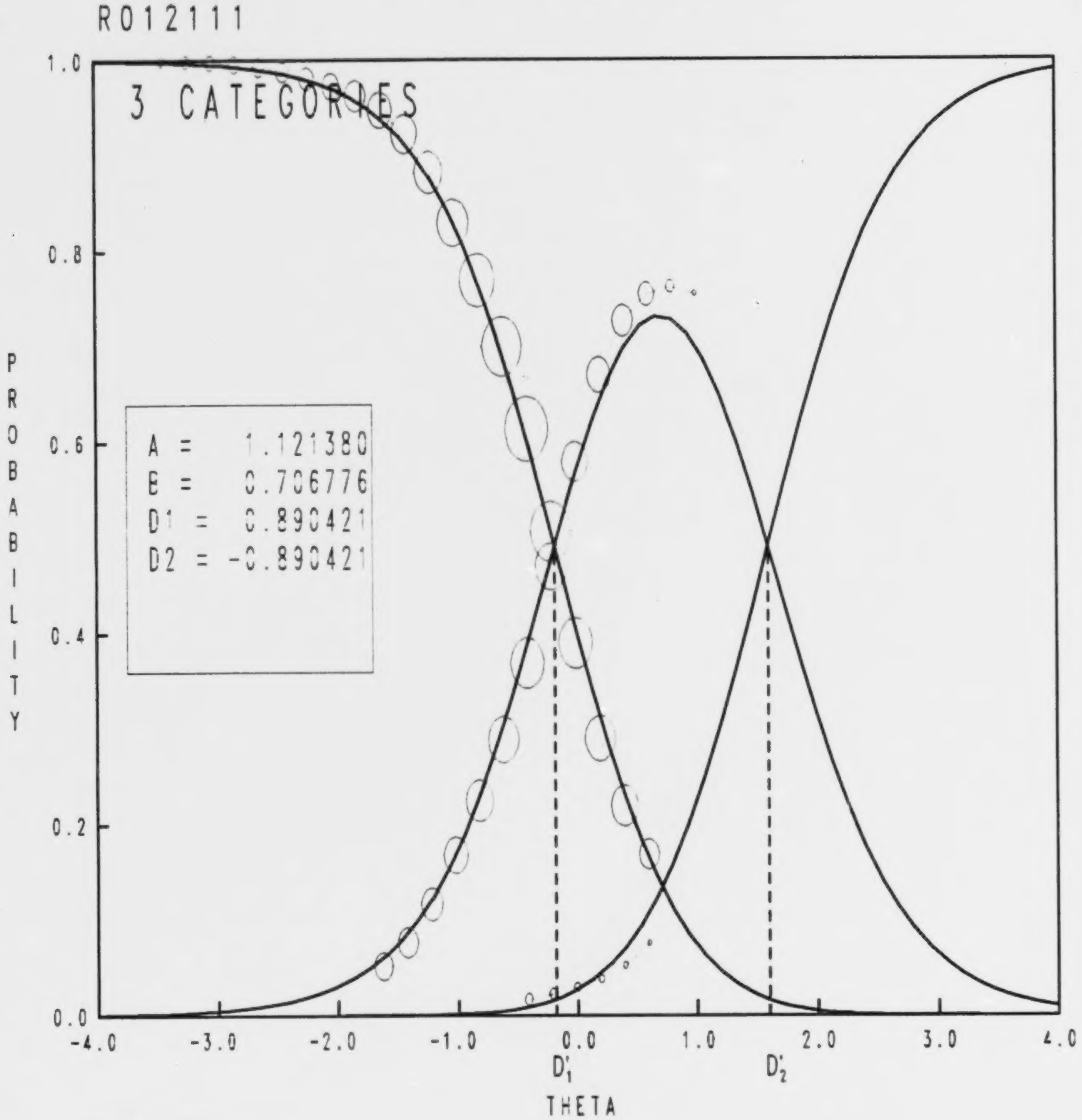


---

* Solid lines represent model-based category traces; ovals represent empirical data.

Figure 12-4

Example Cross-sectional Polytomous Item (R012111, Age 9)
After Collapsing Categories 0 and 1*



---

* Solid lines represent model-based category traces; ovals represent empirical data.

based expected proportions for each age/grade to the single, model-based item response function fit by BILOG/PARSCALE. Items that showed clear evidence of functioning differently across age/grades were treated as separate items for each age/grade level—that is, separate item response functions were estimated for each age/grade level. As was the case with deleting items, in making decisions about scaling items separately by age/grade level, a balance was sought between being too stringent, hence splitting too many items and possibly damaging the vertical link between the age/grades, and being too lenient, hence including items with model fit poor enough to endanger the types of model-based inferences made from NAEP results.

Figure 12-5 shows item R013410 from the Reading to Perform a Task scale, which was presented to both age 13/grade 8 (rectangles) and age 17/grade 12 (diamonds). For proficiency values less than -0.5, the two sets of symbols diverge substantially. Figure 12-6 shows the plots for the item treated separately by age/grade. The remaining misfit at age 17/grade 12 is within acceptable limits. Similarly, Figures 12-7, 12-8, and 12-9 show plots for polytomous item R012607, administered at age 9/grade 4 and age 13/grade 8, from the Reading for Literary Experience scale. At age 9/grade 4, the fit is much better for the split version of the item. At age 13/grade 8, the vertical distances between model prediction line and symbols has decreased, and any remaining lack of fit is not excessive. Table 12-18 lists the items that were treated separately by age/grade level.

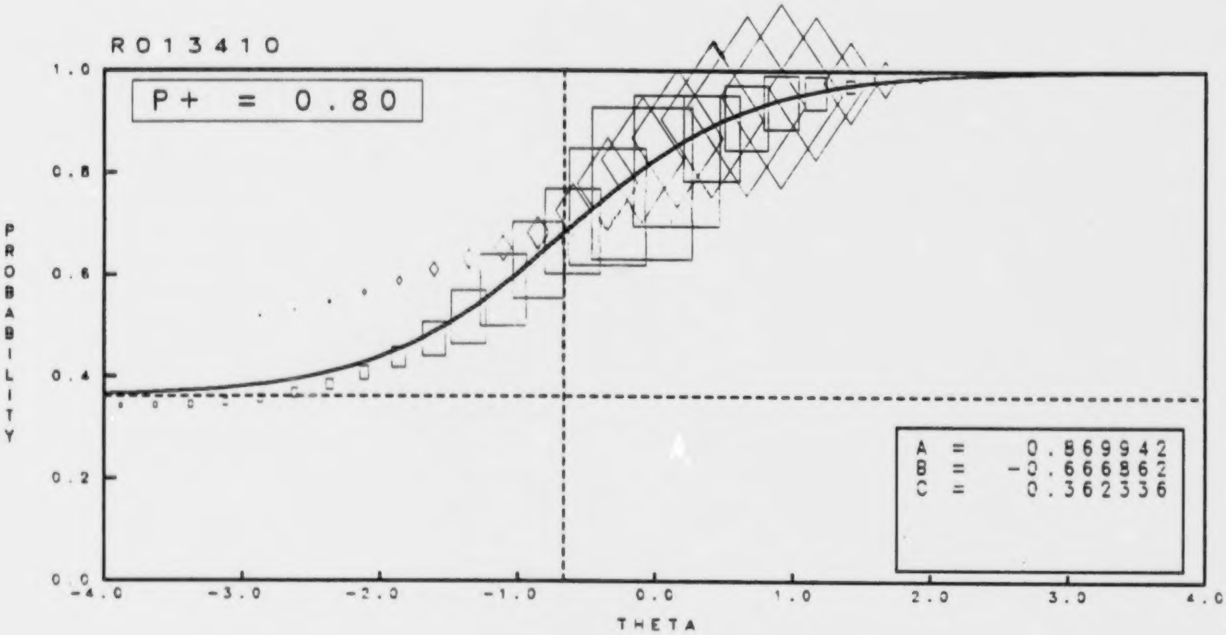
Table 12-18
Items Scaled Separately by Age/Grade

Scale	NAEP ID	Block	Type of Item	Grades Affected
Reading for Literary Experience	R012607	RE	Polytomous	4 and 8
	R012609	RE	Dichotomous	4 and 8
	R013105	RC	Dichotomous	8 and 12
	R013106	RC	Polytomous	8 and 12
Reading to Gain Information	R012702	RG	Dichotomous	4 and 8
	R012705	RG	Dichotomous	4 and 8
	R012708	RG	Polytomous	4 and 8
Reading to Perform a Task	R013304	RI	Dichotomous	8 and 12
	R013305	RI	Dichotomous	8 and 12
	R013410	RJ	Dichotomous	8 and 12

A list of the items scaled for each of the age/grades, along with their final item parameter estimates, appears in Appendix E.

Figure 12-5

Example Cross-sectional Dichotomous Item (R013410) Demonstrating Differential Item Functioning Across Age/grade Samples for Age 13/grade 8 and Age 17/grade 12*

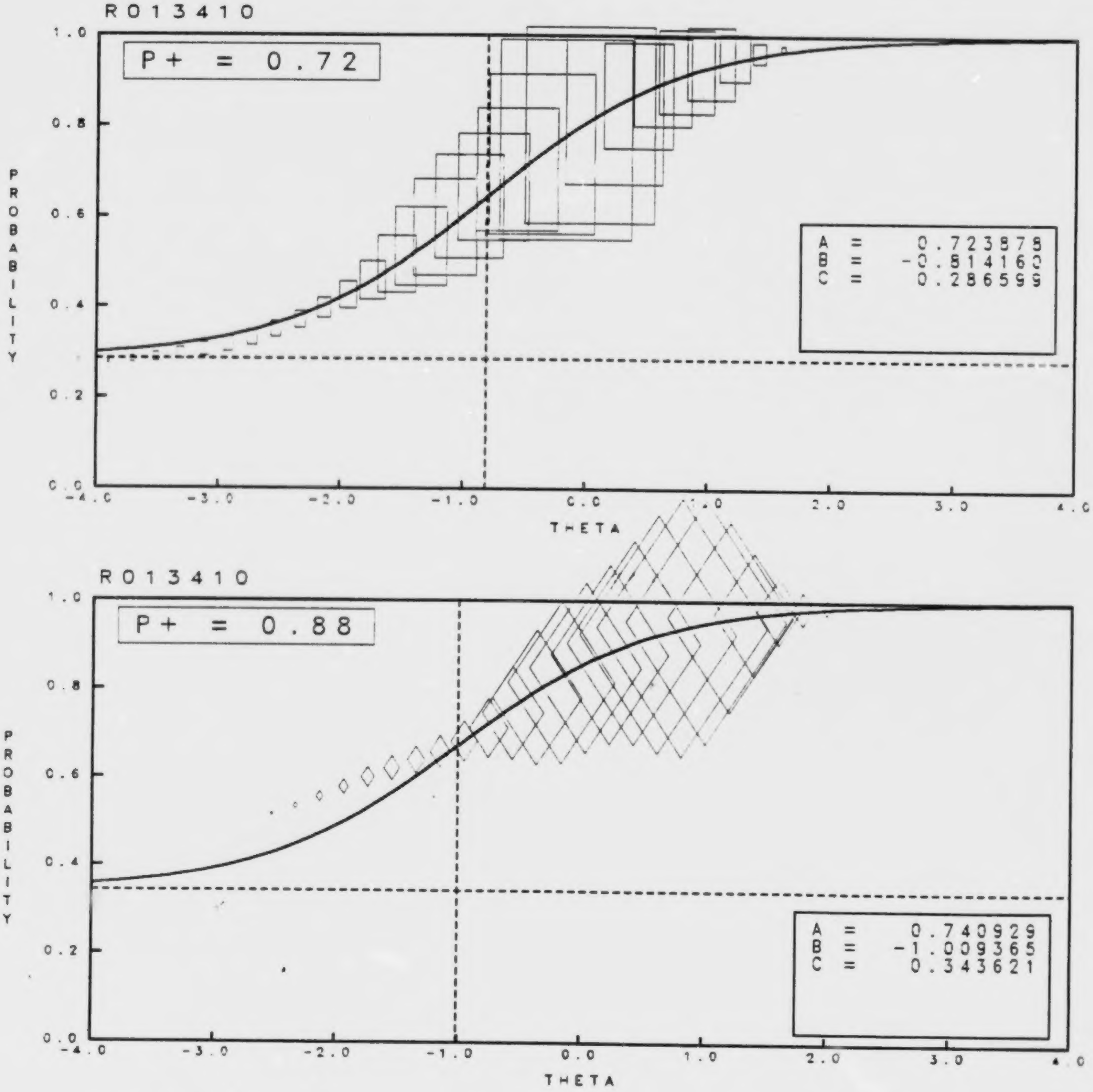


---

* Rectangles represent age 13/grade 8 data; diamonds represent age 17/grade 12 data.

Figure 12-6

Example Cross-sectional Item (R013410)
Fitting Separate Item Response Functions for Each Age/Grade Sample*

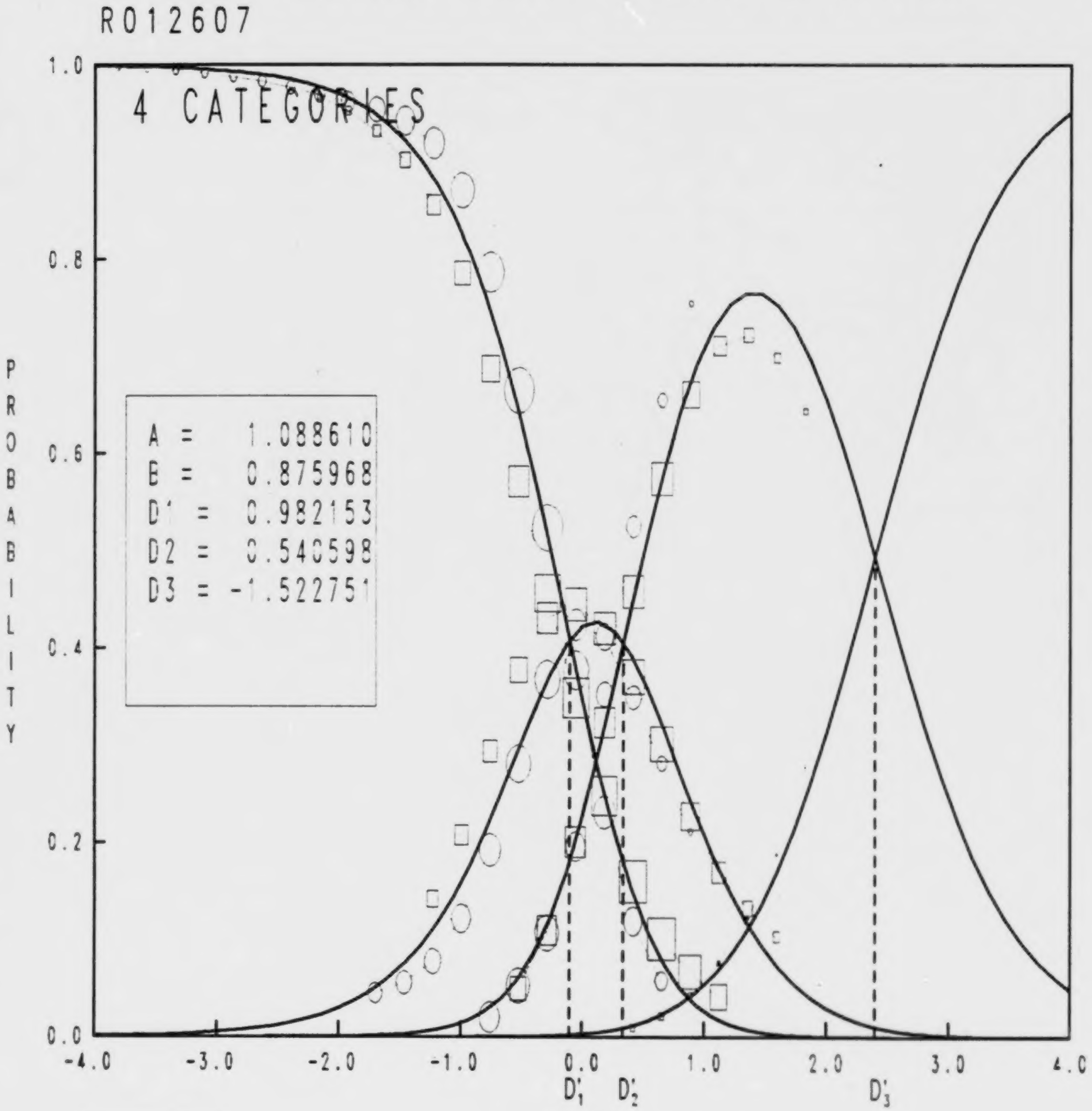


---

* Rectangles represent age 13/grade 8 data; diamonds represent age 17/grade 12 data.

Figure 12-7

Example Cross-sectional Polytomous Item (R012607) Demonstrating Differential Item Functioning Across Age/grade Samples for Age 9/grade 4 and Age 13/grade 8*



---

* Solid line represents common, model-based category traces; ovals represent age 9/grade 4 data; rectangles represent age 13/grade 8 data.

Figure 12-8

Example Cross-sectional Polytomous Item (R012607)
Fitting Separate Item Response Functions for Each Age/grade Sample: Age 9/grade 4

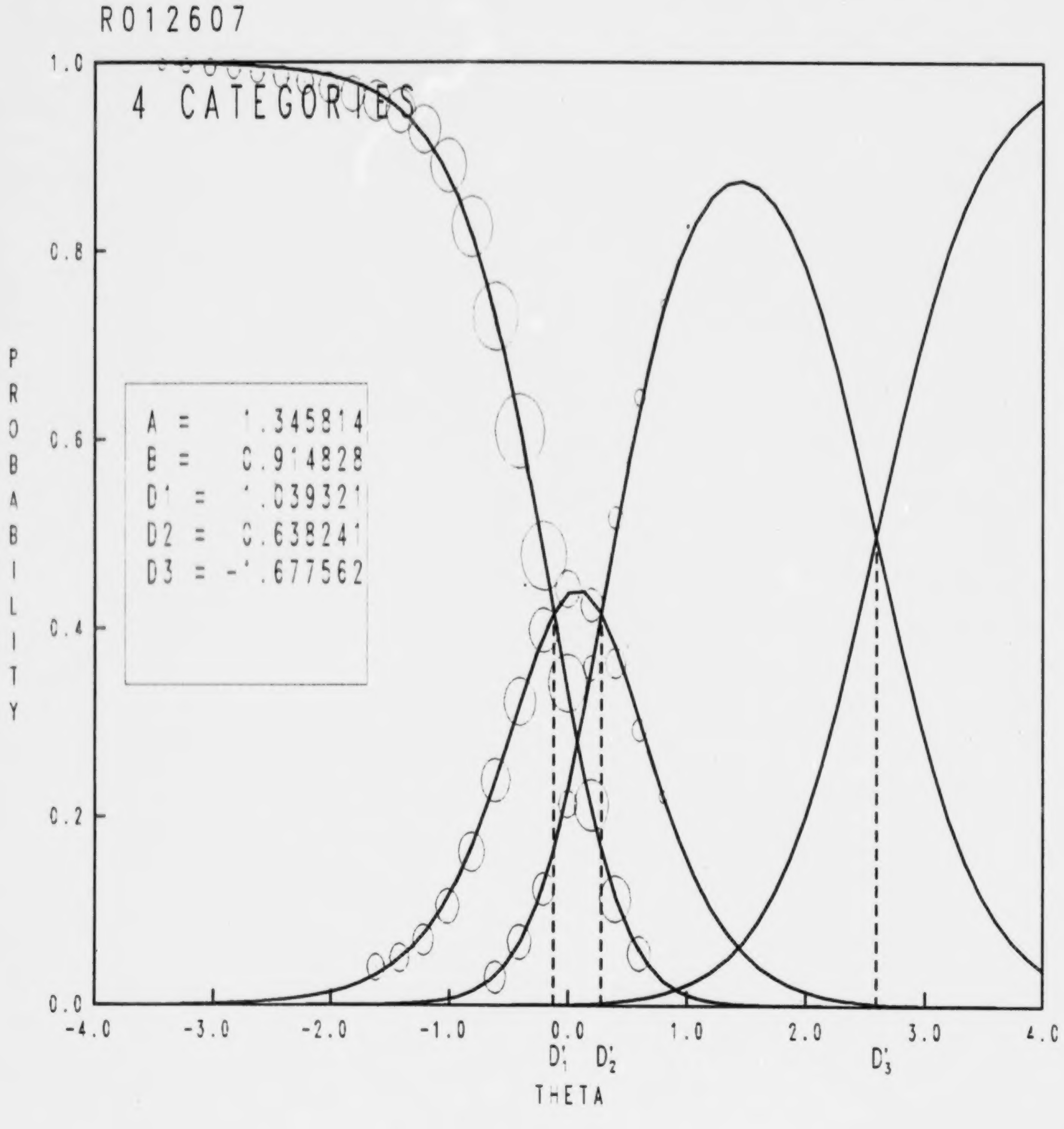
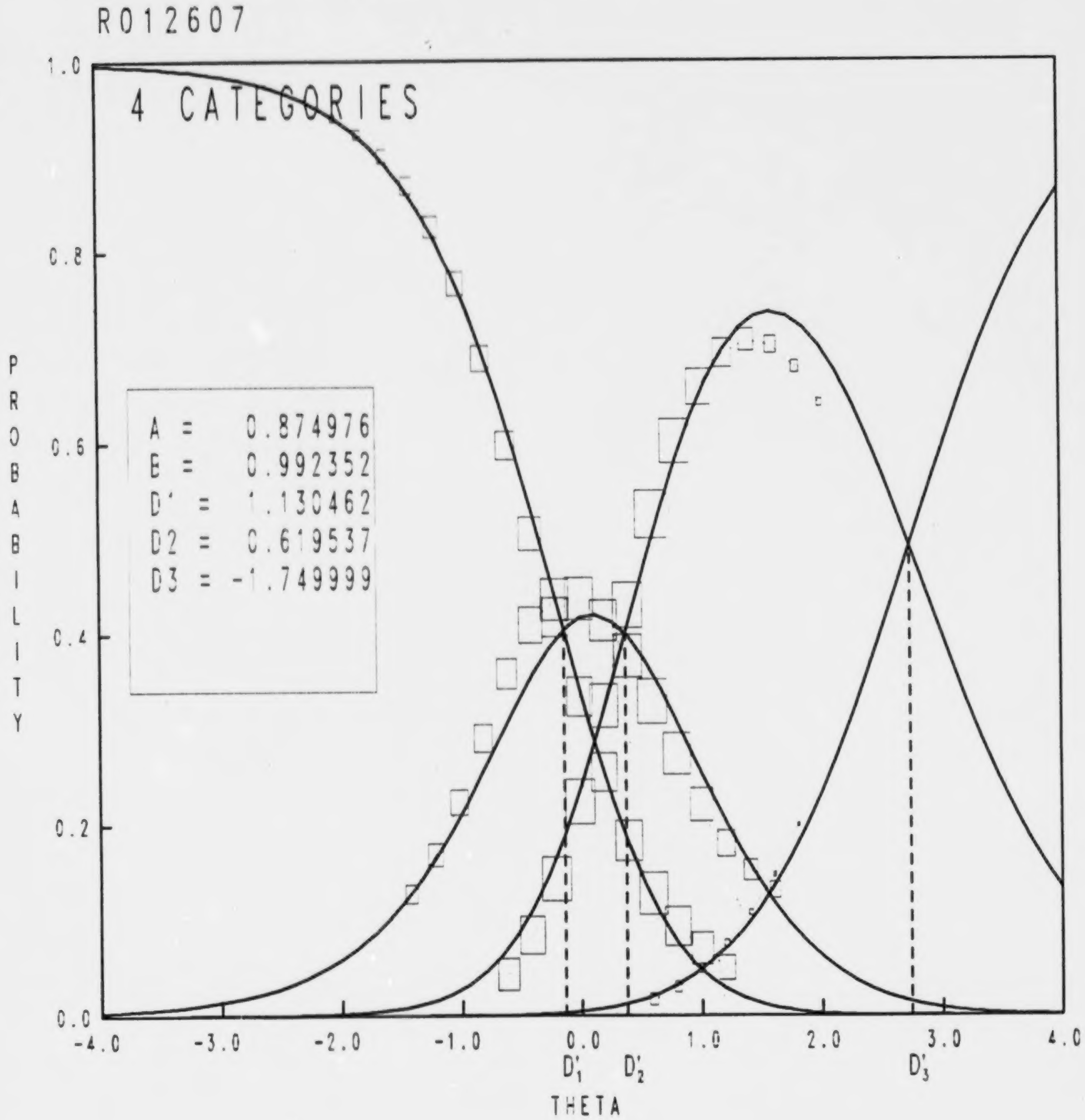


Figure 12-9

Example Cross-sectional Polytomous Item (R012607)
Fitting Separate Item Response Functions for Each Age/Grade Sample: Age 13/grade 8



12.2.5 Generation of Plausible Values

Multivariate plausible values were generated for each age/grade group separately using the CGROUP program described in Chapter 11. Final student weights were used in this analysis. Plans for reporting required analyses examining the relationships between proficiencies and a large number of background variables. The background variables included student demographic characteristics (e.g., race/ethnicity of the student, highest level of education attained by parents), students' perceptions about reading, student behavior both in and out of school (e.g., amount of television watched daily, amount of homework done each day), and a variety of other aspects of the educational, social, and financial environment of the schools they attended. For age 9/grade 4, information was also collected from students' teachers about the types of educational practice to which the students were exposed.

As described in Chapter 11, to avoid bias in reporting results and to minimize biases in secondary analyses, it was desirable to incorporate a large number of independent variables in the conditioning model. When expressed in terms of contrast coded main effects and interactions, the number of variables to be included totaled 368 for age 9/grade 4, 198 for age 13/grade 8, and 218 for age 17/grade 12. The much larger number for age 9/grade 4 reflects the number of contrasts from the teacher questionnaire. Appendix F provides a list of the full set of contrasts defined.

Some of these contrasts involved relatively small numbers of individuals and some were highly correlated with other contrasts or sets of contrasts. Given the large number of contrasts, an effort was made to reduce the dimensionality of the predictor variables. As was done for the 1990 and 1992 mathematics assessments, the original background variable contrasts were standardized and transformed into a set of linearly independent variables by extracting separate sets of principal components at each age/grade level. The principal components, rather than the original variables, were used as the independent variables in the conditioning model. As was done for the mathematics assessment, the number of principal components was the number required to account for at least 90 percent of the variance in the original contrast variables. Research based on data from the 1990 Trial State Assessment in mathematics suggests that results obtained using such a subset of components will differ only slightly from whose obtained using the full set (Mazzeo, Johnson, Bowker, & Fong, 1992). Table 12-19 contains a list of the number of principal components included in conditioning, as well as the proportion of variance accounted for by the conditioning model for each age/grade.

For each age/grade, Table 12-20 provides estimated residual variance for each purpose-of-reading scale and the residual correlation matrix between the reading scales. The values, taken directly from the output of the CGROUP program, are estimates of relationships between the subscales *conditional on the set of principal components included in the conditioning model*. The *marginal* correlations between the purpose-of-reading scales are presented in Table 12-21. As would be expected, they are higher than the conditional correlations.

Table 12-19

Proportion of Proficiency Variance Accounted for by the Conditioning Model for the Reading Main Assessment

Grade	Number of Conditioning Contrasts	Number of Principal Components	Proportion of Proficiency Variance		
			Reading for Literary Experience	Reading to Gain Information	Reading to Perform a Task
4	368	169	.540	.593	—
8	198	111	.487	.540	.566
12	218	115	.575	.582	.580

Table 12-20

Conditional Correlations and Variances from Conditioning (CGROUP)

	Reading for Literary Experience	Reading to Gain Information	Reading to Perform a Task
Grade 4			
Reading for Literary Experience	1.000	—	—
Reading to Gain Information	0.846	1.000	—
Residual Variance	0.284	0.273	—
Grade 8			
Reading for Literary Experience	1.000	—	—
Reading to Gain Information	0.819	1.000	—
Reading to Perform a Task	0.787	0.825	1.000
Residual Variance	0.308	0.260	0.398
Grade 12			
Reading for Literary Experience	1.000	—	
Reading to Gain Information	0.728	1.000	—
Reading to Perform a Task	0.539	0.681	1.000
Residual Variance	0.306	0.175	0.282

Table 12-21

Unconditional Correlations of Reading Scales*

	Reading for Literary Experience	**Reading to Gain Information**	**Reading to Perform a Task**
Grade 4			
Reading for Literary Experience	1.000	—	—
Reading to Gain Information	0.872	1.000	—
Grade 8			
Reading for Literary Experience	1.000	—	—
Reading to Gain Information	0.881	1.000	—
Reading to Perform a Task	0.862	0.893	1.000
Grade 12			
Reading for Literary Experience	1.000	—	
Reading to Gain Information	0.846	1.000	—
Reading to Perform a Task	0.754	0.830	1.000

---

* These correlations are based on the "grade-only" portion of each sample.

12.2.6 The Final Proficiency Scales

12.2.6.1 Purposes-of-reading Scales

Like all IRT scales, the reading scales have a linear indeterminacy that may be resolved by an arbitrary choice of origin and unit-size for each scale. Similar to previous NAEP assessments, the scale mean and standard deviation were set to 250.0[4] and 50.0 using the transformation:

$$\theta_{\text{proficiency}} = A \bullet \theta_{\text{calibrated}} + B.$$

where $\theta_{\text{proficiency}}$ denotes values on the final transformed scale and $\theta_{\text{calibrated}}$ denotes values on the original calibration scale from BILOG/PARSCALE.

For the Reading for Literary Experience and Reading to Gain Information scales, the linear indeterminacy was resolved by transforming the mean and standard deviation of the three age/grade samples combined together to the 250.0, 50.0 metric. The third scale, Reading to Perform a Task, was presented only at age 13/grade 8 and age 17/grade 12. For this scale, the means and standard deviations of each of these two age/grade samples were matched to the average of the other two scales to determine the appropriate transformation. The constants for the linear transformation for each scale are given in Table 12-22.

Table 12-22
Coefficients of Linear Transformations of the Purposes-of-reading Scales
from the Calibrating Scale Units to the Units of Reporting Proficiency

Scale	A	B
Reading for Literary Experience	50.54	248.43
Reading to Gain Information	50.18	257.21
Reading to Perform a Task	42.69	270.88

12.2.6.2 The Composite Proficiency Scales

For the national assessment, a composite scale was created as an overall measure of reading proficiency. The composite was a weighted average of plausible values on the purposes-of-reading scales (Reading for Literary Experience, Reading to Gain Information, and, at age 13/grade 8 and age 17/grade 12, Reading to Perform a Task). The weights for the scales were proportional to the importance assigned to each reading purpose contained in the assessment specifications given in the *Reading Framework*. The weights are given in Table 12-23. As indicated in Table 2-4 of Chapter 2, the weights for each of the purposes of reading are similar to the actual proportion of assessment time devoted to that purpose. In developing the composite scale, the weights were applied to the plausible values for each reading purpose as expressed in terms of the final scale (i.e., after transformation from the provisional BILOG/PARSCALE scales).

---

[4] Previous NAEP scales have used an overall mean of 250.5. Beginning with the 1992 reading assessment, new NAEP scales will use the mean of 250.0.

Table 12-23
Weights Used for Each Purpose-of-reading Scale to Form the Reading Composite

Scale	Age 9/Grade 4	Age 13/Grade 8	Age 17/Grade 12
Reading for Literary Experience	.55	.40	.35
Reading to Gain Information	.45	.40	.45
Reading to Perform a Task	—	.20	.20

Finally, it is necessary to caution that, although the reading composite is expressed in units that seem similar to the long-term trend reading scale, it is not appropriate to compare scores. The transformation chosen to resolve the linear indeterminacy in the reading composite is a convenient transformation, but it is only one of a conceptually infinite number of such transformations that could have been chosen. Any one of these transformations would have provided equivalent information about the relative standings of subgroups in the population. *Because there is no link, real or implied, in the construction of the reading composite and the purpose-of-reading scales to either the mathematics or science assessments or to the previous reading assessments, the comparison of students' reading proficiencies to students' proficiencies in other subject areas is devoid of meaning.*

Overall summary statistics for the composite scale are given in Table 12-24.

Table 12-24
Means and Standard Deviations on the Reading Composite Scale

	All Five Plausible Values	
Grade	Mean	S. D.
4	217.89	36.39
8	260.25	36.01
12	291.05	35.35

12.2.7 Partitioning of the Estimation Error Variance

For each grade, the variance of the final, transformed scale mean was partitioned as described in Chapter 11. This analysis yielded estimates of the proportion of error variance due to sampling students and the proportion due to the latent nature of θ. These estimates are given in Table 12-25 for each purpose-of-reading scale and the composite scale (for stability of the estimates, they are based on 100 plausible values). Additional results, including those by gender and race/ethnicity, are presented in Appendix N.

Table 12-25
Estimation Error Variance and Related Coefficients for the Reading Main Assessment

Grade	Scale	Total Estimation Error Variance	Proportion of Variance Due to...	
			Student Sampling	Latency of θ
4	Reading for Literary Experience	0.90	0.86	0.14
	Reading to Gain Information	1.19	0.89	0.11
	Composite	0.89	0.94	0.06
8	Reading for Literary Experience	0.73	0.79	0.21
	Reading to Gain Information	0.63	0.85	0.15
	Reading to Perform a Task	0.72	0.76	0.24
	Composite	0.56	0.92	0.08
12	Reading for Literary Experience	0.59	0.58	0.42
	Reading to Gain Information	0.35	0.86	0.14
	Reading to Perform a Task	0.43	0.70	0.30
	Composite	0.35	0.84	0.07

Chapter 13

DATA ANALYSIS FOR THE MATHEMATICS ASSESSMENT[1]

Frank Jenkins and Edward M. Kulick

Educational Testing Service

This chapter describes the analyses performed on the responses to the cognitive and background items in the 1992 assessment of mathematics. The emphasis of this chapter is on the methods and results of procedures used to develop the IRT-based scale scores. However, some attention is given to the analysis of constructed-response items. The theoretic underpinnings of the IRT and plausible values methodology described in this chapter are given in Chapter 11.

The objectives of the mathematics analyses were to

- prepare scale values and perform all analyses necessary to produce a long-term trend report in mathematics. The mathematics trend line includes the years 1973, 1978, 1982, 1986, 1990 and 1992.

- prepare scale values for the main analysis of the main focused-BIB mathematics samples. The scaling of mathematics entailed development of several content and skill area scales and an overall composite.

The student samples that were administered mathematics items in the 1992 assessment are shown in Table 13-1. (See Chapters 1 and 3 for descriptions of the target populations and the sample design used for the assessment.) Data from the first three samples (Math-MainP) were used in the main analysis, data from the next three samples (Math-MainT) were used for the special study of estimation and comprehensive problem-solving, data from Math-Calc were used for the calculator bridge study, and data from the rest of the samples (MS-LTTrend) were collected for long-term trend purposes.

The data from the 1992 main focused-BIB samples were scaled separately from the data for the long-term trend samples. Accordingly, the long-term trend and main analyses are presented in separate sections. Section 13.1 pertains to the scaling of the data from the long-term trend samples; section 13.2 contains information about the scaling of the data from the main samples.

---

[1] Data analysis and scaling were performed primarily by Edward Kulick, Drew Bowker, and Fai Fong, with assistance from David Freund, Steven Isham, and Alfred Rogers.

Table 13-1

NAEP 1992 Mathematics Student Samples

Sample	Booklets	Mode	Cohort Assessed	Time of Testing	Age Definition	Modal Grade	Number Assessed
9 [Math-MainP]	1-26	Print	Age 9/grade 4	Winter	CY	4	9,414
13 [Math-MainP]	1-26	Print	Age 13/grade 8	Winter	CY	8	10,291
17 [Math-MainP]	1-26	Print	Age 17/grade 12	Winter	CY	12	9,499
9 [Math-MainT]	27	Tape	Age 9/grade 4	Winter	CY	4	2,054
13 [Math-MainT]	27	Tape	Age 13/grade 8	Winter	CY	8	2,416
17 [Math-MainT]	27	Tape	Age 17/grade 12	Winter	CY	12	2,074
9 [Math-Calc]	28	Print	Age 9/grade 4	Winter	CY	4	2,236
9 [MS-LTTrend]	91-93	Tape	Age 9	Winter	CY	4	7,335
13 [MS-LTTrend]	91-93	Tape	Age 13	Fall	CY	8	5,909
17 [MS-LTTrend]	84-85	Tape	Age 17	Spring	Not CY	11	4,359

LEGEND:

Math	Mathematics
MS	Mathematics and science
Calc	Calculator Bridge
MainP	Main assessment, print administration
MainT	Main assessment, tape administration
LTTrend	Long-term trend: Booklets are identical to long-term trend assessment of 1986.
Print	Printed administration
Tape	Audiotape administration
CY	Calendar year: birthdates in 1982, 1978, and 1974 for ages 9, 13, and 17
Not CY	Age 17 only: birthdates between Oct. 1, 1974 and Sept. 30, 1975

13.1 LONG-TERM TREND DATA ANALYSIS

The long-term trend results reported in *NAEP 1992 Trends in Academic Progress* (Mullis, Dossey, Campbell, Gentile, O'Sullivan, & Latham, 1994) are based on paced-tape administrations and occur at all of the age levels. The samples involved in the analysis are shown as 9[MS-LTTrend], 13[MS-LTTrend], and 17[MS-LTTrend] in Table 13-1. For ages 9 and 13, the long-term trend booklets for these samples contained blocks of mathematics and science items by tape recordings (these booklets also contained reading blocks administered in print form). The age 17 long-term trend booklets contained only mathematics and science blocks, both presented by paced-tape recordings. All students received a block of common background questions, distinct for each age. Subject-area background questions were presented in the cognitive blocks. The booklets for the age 9 and age 13 samples (booklets 91-93) are the same as those used for long-term trend assessments in 1986, 1988, and 1990. The booklets for the age 17 sample (booklets 84-85) are the same as those used for the 1986 and 1990 long-term trend assessments. The booklets and the blocks within those booklets are listed in Table 13-5.

Table 13-2 clarifies the relationships between the 1992 long-term trend samples and samples from previous years. For all ages, the paced-tape bridge to the 1986 samples allow direct comparisons with 1990 samples and with 1986 long-term trend samples. There was also a tape administration in 1988 at ages 9 and 13 that was comparable to the other years. However, a tape administration was not conducted at age 17 in 1988. Instead, a noncomparable paper-based assessment was conducted. Hence, 1988 is not included as a point in the long-term trend reporting. In 1986, the mathematics long-term trend items were scaled with common items from the 1977-78 and 1981-82 assessments. Because the 1972-73 assessment had few items in common with the current assessment, data from that assessment was not scaled using the IRT model but was linked to the trend line by a linear transformation involving the mean proportion correct for common items. The 1992 long-term trend assessment was linked to the 1972-73, 1977-78, and 1981-82 assessments through the 1986 assessment. Information about previous assessment years is available in *Expanding the New Design: The NAEP 1985-86 Technical Report* (Beaton, 1988) and *The NAEP 1990 Technical Report* (Johnson & Allen, 1992).

The steps in the mathematics long-term trend analysis are documented in the following sections. As is usual in NAEP analyses, the first step was to gather item-level and block-level information. Next, the items were calibrated and evaluated for IRT model fit. Derived background variables were calculated for use in conditioning and reporting. Following this, item parameter estimates were used to generate plausible values after conditioning on available background variables. Finally, the scales were placed on the mathematics long-term trend proficiency scale used in previous trend assessments.

Table 13-3 indicates the number of items common across different age combinations. Table 13-4 shows the number of items (scaled in 1992) that were common across assessment years. The 1986, 1990, and 1992 assessments had all items in common. For age 9, the number of items common across assessment years 1978 to 1992 was only 35. For age 13, the overlap across all assessments was 56 items; for age 17, the overlap was 54 items. Item parameters were estimated assuming a univariate scale since the number of items presented to each student was small and there were too few items to estimate several content area scales separately.

Table 13-2

NAEP Mathematics Samples Contributing to 1992 Long-term Trend Results, 1973-1992

Cohort	Year	Sample	Subjects	Time of Testing	Mode of Administration	Age Definition	Modal Grade
Age 9	1973	Main	MS	Winter	Tape	CY	4
	1978	Main	M	Winter	Tape	CY	4
	1982	Main	MCS	Winter	Tape	CY	4
	1986	LTTrend*	MS	Winter	Mixed	CY	4
	1990	LTTrend*	MS	Winter	Mixed	CY	4
	1992	LTTrend*	MS	Winter	Tape	CY	4
Age 13	1973	Main	MS	Fall	Tape	CY	8
	1978	Main	M	Fall	Tape	CY	8
	1982	Main	MCS	Fall	Tape	CY	8
	1986	LTTrend*	MS	Fall	Mixed	CY	8
	1990	LTTrend*	MS	Fall	Mixed	CY	8
	1992	LTTrend*	MS	Fall	Tape	CY	8
Age 17	1973	Main	MS	Spring	Tape	Not CY	11
	1978	Main	M	Spring	Tape	Not CY	11
	1982	Main	MCS	Spring	Tape	Not CY	11
	1986	LTTrend*	MS	Spring	Tape	Not CY	11
	1990	LTTrend*	MS	Spring	Tape	Not CY	11
	1992	LTTrend*	MS	Spring	Tape	Not CY	11

---

* Within an age group these samples had common booklets and constituted a trend line.

LEGEND:

M	Mathematics
MCS	Mathematics, civics, and science
MS	Mathematics and science
Tape	Audiotape administration
Mixed	Mathematics and science administered by audiotape, reading administered by print

Main	Main assessment.
LTTrend	Long-term trend. Booklets are identical to long-term trend assessment of 1986.
CY	Calendar year birthdates.
Not CY	Birthdates between October 1 and September 30.

Table 13-3

Numbers of Scaled Mathematics Long-term Trend Items Common Across Ages

Age	Booklets	Number of Items
9 only	91-93	32
13 only	91-93	30
17 only	84-85	41
9 and 13 only	91-93, 91-93	20
9 and 17 only	91-93, 84-85	0
13 and 17 only	91-93, 84-85	27
9, 13, and 17	91-93, 91-93, 84-85	3
Total		153

Table 13-4

Numbers of Scaled Mathematics Long-term Trend Items Common Across Assessments

Assessment Year	Number of Items
Age 9	
1986, 1990, 1992	55
1982, 1986, 1990, 1992	53
1978, 1986, 1990, 1992	35
1978, 1982, 1986, 1990, 1992	35
Age 13	
1986, 1990, 1992	80
1982, 1986, 1990, 1992	79
1978, 1986, 1990, 1992	56
1978, 1982, 1986, 1990, 1992	56
Age 17	
1986, 1990, 1992	71
1982, 1986, 1990, 1992	65
1978, 1986, 1990, 1992	54
1978, 1982, 1986, 1990, 1992	54

In the first phase of the analysis, standard item statistics were calculated. The results serve as a check for data entry errors and as a reasonableness check against results from previous assessments.

In the second phase, the IRT model was fitted to the data across multiple assessments for each age separately. This puts item parameters and ability estimates on the same scale across years. Note that the same item may have different item parameters for different age groups.

The analysis for an age group was completed by the creation of plausible values through a multiple imputation estimation procedure in which item parameter estimates, student responses, and student background information are combined to produce the most precise possible estimates of student ability. Plausible values from the 1992 assessment were transformed to the scale of the 1986 proficiency measures.

The 1992 plausible values for ages 9, 13, and 17 were used to create proficiency means and jackknifed estimates of standard errors for the whole group and for subgroups. These proficiency means form the final point in the longitudinal mathematics trend from 1973 to 1992.

The specifics of the mathematics long-term trend analysis are documented in the following sections.

13.1.1 Item Analysis for the Long-term Trend Assessment

No problems in coding, formats, or data were detected. The correspondence between blocks, booklets, and samples is given for the regular trend assessment in Table 13-5. Note that common labeling of these blocks across ages does not denote common items. The conventional item analysis, with results displayed in Table 13-6, was performed at the block level on the paced-tape trend data.

Table 13-5
Correspondence Between Samples, Booklets, and Blocks for the Mathematics Long-term Trend

Sample	Booklet	Blocks
Age 9 MS-LTTrend	91	M1
	92	M3
	93	M2
Age 13 MS-LTTrend	91	M1
	92	M3
	93	M2
Age 17 MS-LTTrend	84	M1 M2
	85	M3

Table 13-6

Descriptive Statistics for Item Blocks
Mathematics Long-term Trend Samples

Statistic	Block		
	M1	M2	M3*
Age 9			
Number of scaled items	24	26	5
Number of scaled constructed response items	9	9	0
Unweighted sample size	1991	2194	2050
Average weighted proportion correct	.63	.62	.68
Average weighted r-biserial	.59	.63	.82
Weighted alpha reliability	.81	.86	.46
Age 13			
Number of scaled items	36	36	8
Number of scaled constructed response items	9	8	0
Unweighted sample size	2229	2132	2288
Average weighted proportion correct	.66	.60	.65
Average weighted r-biserial	.56	.57	.66
Weighted alpha reliability	.85	.86	.58
Age 17			
Number of scaled items	33	33	5
Number of scaled constructed response items	10	5	1
Unweighted sample size	2205	2205	2206
Average weighted proportion correct	.64	.65	.54
Average weighted r-biserial	.69	.64	.75
Weighted alpha reliability	.91	.88	.54

---

* This block is mostly calculator items, which were not analyzed.

Note: For the item analysis, students who did not respond to any items in the block were omitted; however, such students were assigned proficiencies in the final database.

Table 13-6 contains the number of items, size of the sample administered the block, mean weighted proportion correct, mean weighted r-biserial, and mean weighted alpha as a measure of reliability for each block. The average values were calculated using examinee weights and the items in the block that were scaled. The 1992 item-level statistics were not very different from those for the 1990 assessment, except that the 1992 data showed a consistent increase in mean weighted proportion correct across the blocks for ages 13 and 17. The percent of examinees not reaching items in the trend blocks was almost always zero because the items were administered with a tape recording to pace response time.

13.1.2 Item Parameter Estimation

The scaling process began with the estimation of item parameters. IRT parameter estimation was accomplished by using the NAEP version (Rogers 1991) of the BILOG/PARSCALE program which combines Mislevy and Bock's (1982) BILOG and Muraki and Bock's (1991) PARSCALE computer programs. Item calibration was performed separately for each of the three age groups, using the total combined data from the 1990 and 1992 assessments. Including the 1990 assessment data assures that item parameters will be similar for adjacent assessments so that year-to-year trends will not be distorted by abrupt changes in calibration. The calibration was performed on the entire sample of students, resulting in a range of about 1,800 to 2,600 examinee responses to each item in each assessment year. The calibration was based on student weights that were rescaled for the 1992 data so that the sum of the weights equaled the unweighted sample size. Also, weights for the 1990 data were restandardized to give equal weight to the two assessment years included in the scaling (see Appendix M). As with the previous assessment, calculator items were excluded from the analysis. Because calculators have changed greatly since the start of the long-term trend, it was judged that calculator questions are no longer comparable across time. These items were left in the assessment, since excluding them would have changed the testing context.

Since parameters for items in blocks M1, M2, and M3 were estimated separately for ages 9, 13, and 17, items administered at more than one age have multiple sets of item parameter estimates. Items were examined for lack of fit with the data. Those that exhibited extreme violation of IRT assumptions (i.e., did not have monotonically increasing item characteristic curves) were deleted from the analysis, as they were in previous assessments. Other items were deleted because they were calculator items, which are not considered part of the regular assessment. These excluded items appear in Tables 13-7, 13-8, and 13-9. As a result of these deletions, 55 items were scaled for age 9, 80 items were scaled for age 13, and 71 items were scaled for age 17. Of the 153 noncalculator items that were part of the assessment, seven items (5 percent) were excluded due to poor fit with the data. A list of the items scaled for each of the ages, along with their item parameter estimates, appears in Appendix E.

13.1.3 Derived Background Variables

In the trend analysis, all derived variables based upon background questions were used both for conditioning and in reporting (to define subgroups). Derived reporting and conditioning variables are described in Appendix B.

Table 13-7
Items Deleted from the Age 9 Mathematics Long-term Trend Analysis

Booklet	Block	Item	Reason for Exclusion
91	M1	N252601	Was deleted in prior assessment
		N262502	Was deleted in prior assessment
92	M3	N268221	Calculator item*
		N276021	Calculator item
		N276022	Calculator item
		N276821	Calculator item
		N276822	Calculator item
		N276823	Calculator item
		N277621	Calculator item
		N277622	Calculator item
		N277623	Calculator item
		N284021	Calculator item
		N284022	Calculator item

---

* Note that all calculator items are deleted from the analysis.

Table 13-8
Items Deleted from the Age 13 Mathematics Long-term Trend Analysis

Booklet	Block	Item	Reason for Exclusion
91	M1	N262502	Was deleted in prior assessment
93	M2	N261601	Was deleted in prior assessment
92	M3	N264521	Calculator item*
		N259921	Calculator item
		N276821	Calculator item
		N276822	Calculator item
		N276823	Calculator item
		N278921	Calculator item
		N278922	Calculator item
		N278923	Calculator item
		N278924	Calculator item
		N278925	Calculator item
		N280621	Calculator item
		N280622	Calculator item
		N280623	Calculator item
		N280624	Calculator item
		N280625	Calculator item
		N280626	Calculator item

---

* Note that all calculator items are deleted from the analysis.

Table 13-9
Items Deleted from the Age 17 Mathematics Long-term Trend Analysis

Booklet	Block	Item	Reason for Exclusion
84	M1	N282801	Was deleted in prior assessment
		N285701	Was deleted in prior assessment
84	M2	N266801	Was deleted in prior assessment
		N255301	Was deleted in prior assessment
85	M3	N259921	Calculator item*
		N264321	Calculator item
		N264521	Calculator item
		N267921	Calculator item
		N276821	Calculator item
		N276822	Calculator item
		N276823	Calculator item
		N278921	Calculator item
		N278922	Calculator item
		N278923	Calculator item
		N278924	Calculator item
		N278925	Calculator item
		N280621	Calculator item
		N280622	Calculator item
		N280623	Calculator item
		N280624	Calculator item
		N280625	Calculator item
		N280626	Calculator item
		N285321	Calculator item

---

* Note that all calculator items are deleted from the analysis.

13.1.4 Generation of Plausible Values

The generation of plausible values was conducted independently for each age group. In this approach we used student background information to condition item responses in order to more accurately estimate student abilities. The univariate conditioning program BGROUP (Rogers, 1991) was used to combine NAEP BILOG/PARSCALE item parameters with weighted item responses and background variables to produce posterior ability estimates called plausible values. As defined in Chapter 11, BGROUP is an enhanced version of the original conditioning program, MGROUP (Sheehan, 1985). *Plausible values are not test scores* in the usual sense, but can be used to provide consistent estimates of population characteristics. There were 45 contrasts in the conditioning model at age 9, 48 at age 13, and 55 at age 17. Appendix F gives the codings for the conditioning variables and the estimated conditioning effects for the three age groups. The estimated conditioning effects in the tables are expressed on the scale of the original calibration (i.e., the theta scale). A check was made on the distributions of the plausible values for each age, including inspection of the whole group and subgroup means and standard deviations.

13.1.5 The Final Proficiency Scale

Since the plausible value (theta) scales have a linear indeterminacy, comparisons with previous assessments will be sensible only if the scale is linearly transformed to a meaningful metric. This indeterminacy was resolved by linking the 1992 scales to previous trend scales. The 1992 data needed to be transformed to compensate for linear changes in the scale due to employing newly estimated item parameters and new BGROUP conditioning parameters in 1992. This was accomplished by first reestimating the 1990 student abilities using 1992 item parameters and 1992 BGROUP parameters. The new 1990 ability estimates were then equated to the old 1990 ability estimates by matching the first two moments (i.e., the mean and standard deviation). The constants for this transformation were then applied to the 1992 data. The transformation equations that resulted are:

Age 9: $\theta_{\text{proficiency}} = 33.57 \bullet \theta_{\text{calibrated}} + 229.11$

Age 13: $\theta_{\text{proficiency}} = 32.24 \bullet \theta_{\text{calibrated}} + 271.57$

Age 17: $\theta_{\text{proficiency}} = 30.64 \bullet \theta_{\text{calibrated}} + 305.44$,

where $\theta_{\text{proficiency}}$ denotes an individual's value on the final transformed scale of the 1992 data and $\theta_{\text{calibrated}}$ denotes an individual's value on the original 1992 theta scale. Overall summary statistics for the long-term trend samples are given in Table 13-10.

Table 13-10
Means and Standard Deviations on the Mathematics Long-term Trend Proficiency Scale

Age	All Five Plausible Values	
	Mean	**S. D.**
9	229.6	33.1
13	273.3	30.9
17	306.7	30.1

13.1.6 Partitioning of the Estimation Error Variance

For each grade's scale, the error variance of the transformed proficiency means was partitioned according to the procedure described in Chapter 11. The variance is partitioned into two parts; the proportion of error variance due to sampling students (sampling variance) and the proportion of error variance due to the fact that proficiency, θ, is a latent variable that is estimated rather than observed. Table 13-11 contains estimates of the total error variance, the proportion of error variance due to sampling students, and the proportion of error variance due to the latent nature of θ (for stability of the estimates, these are based on 100 plausible values). More detailed information is available for gender and race/ethnicity subgroups in Appendix N.

Table 13-11
Estimation Error Variance and Related Coefficients for the Mathematics Long-term Trend Assessment

Age	Total Estimation Error Variance	Proportion of Variance Due to...	
		Student Sampling	**Latency of θ**
9	0.70	0.91	0.09
13	0.70	0.94	0.06
17	0.79	0.93	0.07

13.2 MAIN ASSESSMENT DATA ANALYSIS

The main analysis included three assessments: the main focused-BIB assessment, the grade 4 calculator bridge, and the special mathematics assessment. The main focused-BIB assessment yielded national estimates of mathematics proficiency for the entire population and selected subgroups. The calculator bridge was used to estimate the effect of utilizing a new calculator for grade 4 in the main assessment, and then to correct for that effect. The special mathematics assessment for grades 4, 8, and 12 yielded estimates of how adept students are in responding mathematics items involving two specific skills, estimation and complex problem solving.

13.2.1 The Main Focused-BIB Assessment

The data from the main focused-BIB assessment of mathematics (from samples 9[Math-MainP, 13[Math-MainP], and 17[Math-MainP]) were used for main analyses comparing the levels of mathematics achievement for various subgroups of the 1992 target populations. The main assessment included three student cohorts: students who were either in the fourth grade or 9 years old, students who were either in the eighth grade or 13 years old, and students who were either in the twelfth grade or 17 years old. The birth date ranges for age-eligible students were based on the 1982, 1978, and 1974 calendar years respectively for ages 9, 13 and 17. The larger age/grade cohorts were used for scaling purposes to assure an adequate number of students per item, but only the grade-defined subsample was used for reporting. The sampled students in each of these three cohorts were assessed in the winter. The samples in the main assessment are listed in Table 13-1.

The pool of items used in the 1992 mathematics assessment contained a range of constructed-response and multiple-choice questions measuring performance on sets of objectives (NAEP, 1988). The framework for the objectives is described in Chapter 2. A total of 368 distinct mathematics items addressing these objectives were scaled after some items were dropped (see Table 13-24). The number of items per grade was 155, 183, and 179 respectively for grades 4, 8, and 12. The items were classified into five categories based on their content: numbers and operations; measurement; geometry; data analysis, statistics, and probability; and algebra and functions. See Table 13-12 for numbers of items within content area scales. These five categories of items constituted the scales used in 1992 reporting. The items were partitioned into 13 blocks. The blocks were then assigned to 26 booklets according to a balanced incomplete block (BIB) design. Each booklet contained relatively few items from each of the five categories.

In the main samples, each student was administered a booklet containing three blocks of mathematics cognitive items, a block of background questions common to all booklets for a particular age/grade level, and a block of mathematics-related background questions common to all mathematics booklets for a particular age/grade level. The balanced incomplete block design by which the 13 blocks of mathematics cognitive questions were assigned to the 26 booklets for each age/grade level is detailed in Chapter 4. The 13 blocks were not intended to be parallel measuring instruments. For example, several blocks contained only the items designed for calculator usage, and some blocks contained items for ruler and protractor usage. In addition, the proportion of items sampled from the five categories were not exactly the same among the 13 blocks.

Both age-selected and grade-selected students contributed to the main scaling. However, only the "grade-only" portion of the main focused-BIB mathematics samples contributed to the means and percentages of the main assessment results that are reported in the *NAEP 1992 Mathematics Report Card for the Nation and the States* (Mullis, Dossey, Owen, & Phillips, 1993). The 368 unique items were constructed according to several formats, some traditional and some more innovative: regular multiple-choice, "clustered" multiple-choice, short constructed-response, and extended constructed-response. The regular multiple-choice items conformed to the familiar format of a stem followed by several possible answers, with only one answer being correct. Cluster items were derived from a set of three to five multiple-choice questions that

Table 13-12

Main Focused-BIB: Numbers of Scaled Mathematics Items Common Across Grade Levels, by Content Area Scale*

Grade(s)	**Numbers and Operations**	**Measurement**	**Geometry**	**Data Analysis, Statistics & Probability**	**Algebra & Functions**	**Total**
4 only	32	15	11	11	7	76
8 only	17	9	13	7	12	58
12 only	22	16	21**	14	36	109
4 and 8	20	11	12	6	6	55
8 and 12	10	9	7	12	8	46
4 and 12	0	0	0	0	0	0
4, 8, and 12	11	3	4	3	3	24
Total	112	63	68	53	72	368

---

* Totals in this table are not the same as totals in Table 2-8 because in this analysis, some items were deleted, some items were clustered, and estimation items were not included.

** Note that an element from a cluster item was added to the set as a separate item, and was retained for the entire analysis. This increased the total number of items by one.

related to the same basic stem. The cluster score was the number of constituent cluster items that the respondent answered correctly. Short constructed-response items were questions that required an open-ended response (e.g., explaining why the previous question was answered as it was). These question were read by raters who determined whether or not the response was correct. Extended constructed-response items were open-ended questions that required several stages of reasoning or problem-solving. They also were read by raters but were given a score reflecting degree of correctness rather than simply judged right or wrong. As Table 13-13 indicates, of the 155 items at grade 4, 94 were multiple-choice items, 2 were cluster items, 54 were short constructed-response items and 5 were extended constructed-response items. Of 183 items at grade 8, 116 were multiple-choice, 2 were cluster, 59 were short constructed-response and 6 were extended constructed-response. Of the 179 items at grade 12, there were 113 multiple-choice items, 2 cluster items, 58 short constructed-response items and 6 extended constructed-response items.

Table 13-13
Main Focused-BIB: Numbers of Items in Various Formats, by Grade

Grade	Multiple-choice	Regular Constructed-Response	Extended Constructed-Response	Cluster	Total
4	94	54	5	2	155
8	116	59	6	2	183
12	113	58	6	2	179

13.2.2 The Calculator Bridge Assessment for Grade 4

An indication of the effect of using a new type of calculator in the main grade 4 assessment is given in Table 13-14, which lists the average percent correct for blocks in the calculator bridge assessment and comparable blocks in the main assessment. Although the noncalculator block showed a small difference, a larger difference is indicated for the calculator blocks. As a result, there had to be an adjustment for the calculator effect in the analysis. The idea was to put the 1992 main assessment and the calculator bridge on the same scale (assuming they came from the same population) so that any difference between the performance of students who used the new calculator and the performance of students who used the old calculator would be reflected in different item difficulties, rather than in different estimated student abilities. In this way, the main assessment would be "adjusted for" the effect of using a new calculator. The calibration also included the 1990 main sample. The block design for the calibration is given in Table 13-15. The samples are put on the same scale during calibration by a combination of common item and common population equating. Specifically, the IRT calibration was designed according to a model that satisfied the following conditions:

1) The 1992 main and calculator bridge data were sampled from the same population. This was stipulated by setting the prior ability distributions for the samples to be equal.

Table 13-14

Difference in Mean Proportion Correct for Blocks in the Grade 4 Calculator Bridge and the 1992 Main Assessment

Block	Calculator Bridge (Old Calculator)	1992 Main Assessment (New Calculator)	Difference
M3 (noncalculator)*	.4829	.4778	-.0051
M8 (calculator)	.5923	.5662	-.0261
M14 (calculator)	.4430	.4257	-.0173

---

* Block M3 is the noncalculator linking block.

Table 13-15

Blocks Used in the Scaling of the 1992 Mathematics Assessment

Sample	M3	M4	M5	M6	M7	M8*	M9	M10	M11	M12*	M13	M14*	M15	90A	90B	90C
1992 main	X	X	X	X	X	X	X	X	X	X	X	X	X			
Calculator bridge	X					X						X				
1990 main		X	X	X		X								X	X	X

"X" designates an existing block; blank cells are blocks not in the design.

* Calculator block

Note: 90A, 90B, and 90C stand for the 1990 BIB blocks that do not overlap the 1992 assessment. 1990 blocks M5 and M6 reversed names in 1992.

2) The 1992 main and calculator bridge assessments shared some common items that did not require use of a calculator. This was operationalized by setting the item parameters to be identical across samples for block M3.

3) The calculator items for the 1992 main and calculator bridge assessments required the use of different types of calculators and may have different item difficulties. This was stipulated by allowing items in blocks M8 and M14 to be different across the two assessments.

4) The 1992 main and 1990 main assessments have some items in common; that is, items in blocks M4, M5, and M6 are identical across assessments.

5) The calculator bridge and the 1990 main assessments have some items in common; that is, items in block M8 are the same across assessments.

6) The 1990 main sample is from a different population than the other samples; that is, the prior ability distribution for the 1990 main sample is different from the prior ability distribution for the other two samples.

When the IRT calibration was run with items and prior proficiency distributions defined according to conditions 1 through 6 above, estimates from the 1992 main assessment were corrected for the effect of using a new calculator.

13.2.3 Special Mathematics Assessment

The special samples 9[Math-MainT], 13[Math-MainT], and 17[Math-MainT] were taken from the same populations as the main mathematics assessment in order to focus on students' abilities in estimation and complex problem-solving skills.

Each age/grade sample of between 1,500 and 1,800 students was administered one booklet of four cognitive blocks appropriate to the age/grade level. Every booklet had one nonpaced cognitive block in common with the main assessment and a 45-minute audiotaped assessment consisting of three blocks of items—one estimation block and two complex problem-solving skills blocks. In addition, each booklet contained three blocks of noncognitive questions: a block of background questions common to all main focused-BIB booklets for a particular age/grade, a block of mathematics-related background questions common to all main focused-BIB mathematics booklets for a particular age/grade, and a block of three items concerning how difficult the student found the assessment. The number of estimation and complex problem-solving items for each age/grade booklet is shown in Table 13-16.

Table 13-16
Number of Cognitive Items in the 1992 Special Mathematics Assessment

Age/Grade	Estimation Block*	Complex Problem Solving Blocks	
	M16	M17	M18
Age 9/Grade 4	20	10	10
Age 13/Grade 8	22	24	6
Age 17/Grade 12	22	30	12

---

* Age 13/Grade 8 and Age 17/Grade 12 received identical estimation blocks.

Although the estimation items were classified into the five content area scales found in the main assessment, the number of items in a single scale was too small to support stable linking to the main assessment separately by content area scale. Therefore the estimation items were treated as if they constituted a single unidimensional scale.

In the 1990 assessment it was observed that complex problem-solving items (which were called HOTS [higher-order thinking skills] items) had distributions that were very different from items in the main assessment. This observation was also found to be true in the 1992 assessment. For complex problem-solving items there were higher numbers of omits than for regular items, especially for Black and Hispanic students. For example, the omit rate for complex problem-solving items ranged from 1 to 9 percent for White twelfth graders, but ranged from 1 to 26 percent for Black twelfth graders. Another problem with the response distribution was the possible floor effect on ability estimates caused by high numbers of minorities who answered none of the complex problem-solving items correctly. As shown in Table 13-17, at grade 12, 4 percent of the White students answered all of the complex problem-solving items incorrectly, while 8 percent of the Hispanic and 11 percent of the Black students answered them all incorrectly. With so few minority students getting any of the items correct there is a floor effect in that there is little information in the data to distinguish among the complex problem-solving abilities for many of the students. Because of the high omit rates and low number of students getting any correct responses among minority students, it was decided that the complex problem-solving items would be treated in a manner similar to the way the comparable higher-order thinking skills items were analyzed in the 1990 assessment. As a result, the complex problem-solving items were *not* combined with estimation items, and the IRT-based analysis was not applied to these items. Instead, they underwent only a classic item analysis consisting of estimates such as average proportion correct and item/total-test statistics.

The number of overlapping items for the age/grade levels are listed in Table 13-18. Numbers of items in the estimation scale by block and by booklet are given in Appendix E. The tabled values reflect only those items included in the final analysis.

Table 13-17

Percentage of Students with All Items Incorrect or Omitted for Complex Problem-solving Blocks, by Grade and Race/Ethnicity

Grade	Race/Ethnicity	Percentage with All Items Incorrect	
		Block M17	Block M18
4	White Hispanic Black	1 1 1	5 11 16
8	White Hispanic Black	0 0 1	8 17 28
12	White Hispanic Black	1 2 1	4 8 11

Table 13-18

Number of Estimation and Complex Problem-solving Items Common Across Grades

Grade	Number of Items	
	Estimation	Complex Problem-solving
4 only	10	13
8 only	0	0
12 only	0	12
4 and 8 only	0	0
4 and 12 only	0	0
8 and 12 only	12	23
4, 8, and 12	10	7
Total	32	55

13.2.4 Item Analysis

This section contains a detailed description of the item analysis performed using the main focused-BIB sample data. As with the trend assessment the analysis examines items within blocks. In preparation for this step, extended constructed-response and cluster items were polytomously scored, short constructed-response items were dichotomously scored, and derived background variables were calculated. Item statistics such as mean percent correct, average score, item to total score correlations and percent responding in each item category were calculated.

Tables 13-19, 13-20, and 13-21 show the number of items, mean proportion correct, mean item to total score correlation, and alpha reliability for each block administered at each age/grade level for the main assessment. These values were calculated within block only for those items used in the scaling process. The table also gives the number of students who were administered the block and the percent not reaching the last item in the block. These numbers include the students in the grade-only portion of the samples that contributed to the summary statistics provided in the *Mathematics Report Card*. Student weights were used, except for the sample sizes. The results for the blocks administered to each age/grade level indicated that the blocks differ in number of items, average difficulty, reliability, and percent not reaching the last item.

A preliminary item analysis was run before the nonresponse correction to the sampling weights was available, to provide timely guidance for the scaling process. The final item analysis was run after scaling was completed.

As described in Chapter 9, in NAEP analyses (both conventional and IRT-based), a distinction is made between missing responses at the end of each block (not-reached) and missing responses prior to the last completed response (omitted). Not-reached items are those occurring after the last item the student completed in a block. Items that were not reached are treated as if they had not been presented to the examinee, while omitted items are regarded as incorrect. The proportion of students attempting the last item of a block (or, equivalently, 1 minus the proportion not reaching the last item) is often used as an index of the degree of speededness of the block of items.

Standard practice at ETS is to treat all students who did not respond to the last item as if they had not reached that item. For multiple-choice and short constructed-response items, this convention produced a reasonable pattern of results, in that the proportion of students reaching the last item does not differ markedly from the proportion attempting the next-to-last item. However, for the blocks that ended with extended constructed-response items, this convention resulted in an implausibly large drop in the number of students attempting the final item. Therefore, for blocks that ended with an extended constructed-response item, students who attempted the next-to-last item but did not respond to the last item were classified as having intentionally omitted that item.

Tables 13-19, 13-20, and 13-21 contain information about the effect of the position of blocks within booklets on the average percent correct for items within each block presented to the focused-BIB samples for each grade. The averages for the grade-only portion of the

Table 13-19

Descriptive Statistics for Item Blocks by Position Within Booklet and Over All Occurrences
Grade 4

Statistic	Block Position	Block												
		M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	M14	M15
Number of scaled items														
Total		13	14	17	11	10	15	10	6	16	12	12	9	10
Constructed-response		4	0	4	0	3	1	2	6	5	7	5	2	3
Extended constructed-response		0	0	0	0	1	0	1	0	0	0	1	1	1
Unweighted sample size	1	567	524	538	555	531	557	555	549	546	563	573	567	542
	2	555	535	557	531	543	549	542	563	549	557	548	569	568
	3	556	562	548	551	549	539	533	564	531	549	568	545	564
	ALL	1678	1621	1643	1637	1623	1645	1630	1676	1626	1669	1689	1681	1674
Average weighted proportion correct	1	.47	.44	.43	.41	.40	.58	.45	.38	.49	.49	.51	.42	.48
	2	.48	.43	.43	.41	.39	.58	.45	.37	.50	.49	.52	.43	.47
	3	.49	.43	.42	.41	.40	.55	.42	.35	.49	.49	.52	.45	.47
	ALL	.48	.43	.43	.41	.40	.57	.44	.37	.49	.49	.51	.43	.47
Average weighted r-biserial	1	.58	.50	.54	.67	.57	.54	.60	.80	.59	.60	.66	.58	.62
	2	.62	.50	.57	.63	.58	.54	.63	.84	.57	.63	.64	.63	.65
	3	.60	.53	.57	.69	.59	.54	.61	.85	.60	.63	.64	.63	.64
	ALL	.60	.51	.56	.66	.58	.54	.61	.83	.59	.62	.65	.61	.64
Weighted alpha reliability	1	.65	.56	.70	.72	.53	.70	.62	.62	.75	.67	.69	.54	.62
	2	.70	.55	.74	.68	.55	.69	.66	.65	.73	.72	.67	.62	.66
	3	.67	.63	.72	.74	.56	.66	.63	.67	.76	.70	.69	.62	.65
	ALL	.67	.58	.72	.72	.55	.68	.64	.65	.76	.70	.68	.60	.64
Weighted proportion of students reaching the last item	1	.82	.94	.67	.76	.91	.77	.99	.89	.89	.66	.94	.90	.99
	2	.87	.96	.80	.77	.91	.84	.99	.90	.90	.69	.96	.93	.98
	3	.82	.94	.79	.78	.93	.89	.99	.92	.95	.67	.95	.93	.99
	ALL	.84	.95	.75	.77	.92	.84	.99	.90	.91	.67	.95	.92	.98

Table 13-20

Descriptive Statistics for Item Blocks by Position Within Booklet and Over All Occurrences
Grade 8

Statistic	Block Position	Block												
		M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	M14	M15
Number of scaled items														
Total		13	21	21	16	13	18	9	7	19	9	11	9	17
Constructed-response		3	0	5	16	5	2	3	7	6	2	4	2	4
Extended constructed-response		1	0	0	0	1	0	1	0	0	1	1	1	0
Unweighted sample size	1	579	584	585	605	588	568	587	589	588	585	611	606	583
	2	591	569	576	598	587	603	585	573	587	585	580	600	607
	3	576	596	593	586	584	580	585	596	599	584	609	593	573
	ALL	1746	1749	1754	1789	1759	1751	1757	1758	1774	1754	1800	1799	1763
Average weighted proportion correct	1	.48	.58	.69	.55	.62	.49	.35	.53	.63	.40	.50	.34	.56
	2	.51	.59	.67	.56	.61	.49	.35	.54	.63	.42	.51	.36	.55
	3	.50	.57	.67	.57	.62	.49	.34	.50	.62	.41	.49	.36	.56
	ALL	.50	.58	.68	.56	.62	.49	.35	.52	.63	.41	.50	.35	.56
Average weighted r-biserial	1	.66	.54	.69	.70	.67	.59	.61	.78	.60	.64	.63	.61	.61
	2	.69	.55	.70	.66	.65	.62	.59	.79	.62	.65	.61	.55	.60
	3	.67	.51	.67	.70	.63	.61	.63	.80	.66	.67	.62	.56	.57
	ALL	.67	.53	.68	.69	.65	.61	.61	.79	.63	.66	.62	.57	.60
Weighted alpha reliability	1	.75	.74	.86	.84	.76	.76	.58	.59	.76	.57	.66	.58	.73
	2	.77	.76	.87	.81	.75	.77	.57	.63	.79	.60	.65	.48	.73
	3	.76	.71	.86	.84	.72	.77	.62	.65	.81	.62	.66	.49	.70
	ALL	.76	.74	.87	.83	.74	.77	.59	.63	.79	.59	.66	.52	.72
Weighted proportion of students reaching the last item	1	.96	.90	.86	.92	.97	.68	.96	.93	.90	.98	.98	.92	.87
	2	.98	.89	.81	.91	.98	.70	.94	.94	.90	.98	.98	.92	.82
	3	.96	.91	.87	.93	.98	.74	.96	.92	.92	.98	.98	.97	.85
	ALL	.97	.90	.85	.92	.97	.71	.95	.93	.90	.98	.98	.94	.84

Table 13-21

Descriptive Statistics for Item Blocks by Position Within Booklet and Over All Occurrences
Grade 12

Statistic	Block Position	Block												
		M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	M14	M15
Number of scaled items														
Total		14	22	19	17	17	21	9	10	14	9	9	10	12
Constructed-response		4	0	3	17	6	4	2	6	3	3	5	3	2
Extended constructed-response		0	0	0	0	0	0	1	1	0	1	1	1	1
Unweighted sample size	1	524	540	551	541	555	534	520	534	551	539	516	525	536
	2	526	552	523	540	562	527	544	523	525	539	549	520	521
	3	530	519	539	527	538	542	554	537	541	538	506	537	533
	ALL	1580	1611	1613	1608	1655	1603	1618	1594	1617	1616	1571	1582	1590
Average weighted proportion correct	1	.40	.66	.59	.55	.41	.53	.47	.43	.45	.43	.41	.43	.44
	2	.41	.69	.58	.56	.42	.53	.47	.45	.45	.45	.41	.46	.44
	3	.39	.68	.57	.56	.42	.52	.49	.44	.44	.45	.42	.45	.45
	ALL	.40	.67	.58	.55	.42	.53	.47	.44	.45	.44	.41	.45	.44
Average weighted r-biserial	1	.67	.59	.61	.71	.68	.64	.60	.62	.59	.73	.68	.57	.56
	2	.70	.54	.59	.73	.67	.64	.60	.63	.60	.74	.68	.64	.57
	3	.69	.55	.59	.73	.66	.60	.61	.64	.63	.76	.67	.63	.55
	ALL	.69	.56	.59	.72	.67	.63	.60	.63	.61	.75	.67	.61	.56
Weighted alpha reliability	1	.72	.77	.78	.82	.71	.82	.56	.58	.68	.62	.65	.49	.60
	2	.74	.72	.77	.84	.71	.81	.58	.62	.68	.65	.64	.60	.59
	3	.72	.73	.76	.84	.70	.77	.58	.60	.72	.66	.64	.60	.60
	ALL	.73	.74	.77	.83	.71	.80	.57	.60	.69	.64	.64	.57	.60
Weighted proportion of students reaching the last item	1	.72	.86	.58	.74	.66	.49	.97	.91	.70	.97	.92	.95	.99
	2	.70	.82	.62	.77	.65	.47	.93	.89	.73	.96	.95	.92	.98
	3	.68	.86	.59	.68	.69	.56	.93	.93	.71	.94	.93	.95	1.00
	ALL	.70	.85	.60	.73	.67	.51	.94	.91	.71	.96	.93	.94	.99

focused-BIB samples show that the order of blocks within booklets did not have a large or consistent effect on proficiency in the mathematics focused-BIB assessment.

Table 13-22 shows estimates for special sample items that parallel the statistics described above for the main assessment. These blocks are presented separately because they were not analyzed together with other scales. Moreover, the results from the special sample items were not part of the composite scale results.

Table 13-22
Descriptive Statistics for the Estimation Block (Block M16 or MP), by Grade

Statistic	Grade 4	Grade 8	Grade 12
Number of scaled items	20	22	22
Number of scaled constructed response items	0	0	0
Unweighted sample size	1562	1769	1526
Average weighted item score	.56	.56	.66
Average weighted r-polyserial	.41	.47	.54
Weighted alpha reliability	.51	.68	.74
Weighted proportion of students attempting last item	1.00	1.00	1.00

13.2.5 Constructed-response Items

As indicated in Table 13-13, about one-third of the mathematics items were constructed-response. Short constructed-response items were given a right/wrong scoring. The categories of responses for the items and the number of responses that were rescored for each item are indicated in Appendix K. The percent agreement for the raters and Cohen's Kappa, a reliability estimate appropriate for items that are dichotomized, are also given in the tables. The sample sizes listed in the tables correspond to the samples used in calculating the rater reliability.

In general, the rater reliability of the scoring for dichotomized responses was quite high. Cohen's Kappa reliabilities ranged over items from 0.77 to 0.99 for age 9, from 0.74 to 0.99 for age 13, and from 0.79 to 0.99 for age 17.

Chapter 7 discusses the definition of the item ratings and describes the process by which teams of raters scored the constructed-response items. This discussion includes the rating definitions for short and extended constructed-response items as well as the range of interrater reliabilities that occurred. Extended constructed-response items were scored on a scale from 1 to 5 to reflect degrees of knowledge. Rating information on extended constructed-response items can be found in Appendix K, which lists the sample sizes, percent agreement, and Cohen's Kappa reliability index.

13.2.6 Differential Item Function Analysis

A differential item functioning (DIF) analysis of the main-assessment mathematics items was done in order to guide committees in identifying biased items. Sample sizes were large enough to compare male and female students, White and Black students, and White and Hispanic students using a modification of the Mantel-Haenszel procedure as adapted by Holland and Thayer (1988), which is described in Chapter 9. The purpose of the analysis is to identify items that should be examined more closely for possible bias.

Currently, the Mantel-Haenszel procedure may only be used for dichotomous data. Thus, for the purposes of DIF analyses *only*, the extended constructed-response items were dichotomized; unsatisfactory and partial responses were treated as incorrect, and essential and extensive responses were treated as correct. The "grade-only" portion of the main focused-BIB was used for DIF analyses. Weights were rescaled separately for each comparison, as described in Chapter 9. DIF analyses were conducted separately by grade. The DIF index generated by the Mantel-Haenszel procedure is commonly used to ETS to place items into one of three categories: A, B, or C. "A" items exhibit no DIF, while "C" items exhibit a strong indication of DIF and should be examined more closely. Positive values of the index indicate items that are differentially easier for the "focal" group (female, Black, or Hispanic students) than for the "reference" groups (male or White students). Similarly, negative values indicate items that are differentially harder for the focal group than the reference group. Table 13-23 shows the distribution of items by five categories of DIF. "B+" and "C+" items were easier for the focal group, while "B-" and "C-" items were harder for the focal group.

Each item was subjected to at least three DIF analyses (more if an item was used across grades). For grade 4, 176 items were analyzed; there were only 12 occurrences of significant DIF (the "C" category), evidenced by 11 separate items. Of 205 items at grade 8, there were 17 occurrences of significant DIF, by 17 different items. At grade 12, where 201 items were analyzed, there were 10 occurrences of significant DIF, by 9 different items.

Following standard practice at ETS for DIF analyses conducted on final forms, all "C" items were reviewed by a committee of trained test developers and subject-matter specialists. As described in Chapter 9, such committees are charged with making judgments about whether or not the differential difficulty of an item is *unfairly* related to group membership (bias). The committee assembled to review NAEP items included both ETS staff and outside members with expertise in the field. It was the committee's judgment that none of the "C" items for the national or the Trial State Assessment data were functioning differentially due to factors irrelevant to test objectives; that is, none were biased. Hence, none of the items were removed from scaling due to differential item functioning. It should be noted that if all differentially functioning items were automatically deleted without an independent assessment of potential for bias, the result would be to obscure differences between subgroups, which would have potentially important policy implications.

Table 13-23
Mathematics DIF Category by Grade

Grade	DIF Category*	Analysis		
		Male/Female	**White/Black**	**White/Hispanic**
4	C+	1	3	0
	B+	14	9	7
	A	147	143	154
	B-	12	16	14
	C-	2	5	1
8	C+	2	4	0
	B+	13	11	7
	A	171	169	190
	B-	15	14	8
	C-	4	7	0
12	C+	3	2	1
	B+	15	16	12
	A	169	161	176
	B-	13	14	8
	C-	1	3	0

---

* A = no indication of DIF; B = Weak indication of DIF; C = strong indication of DIF.

13.2.7 Estimation of Item Parameters

The BILOG/PARSCALE computer program was used to estimate the item parameters for the main assessment and for the special estimation study. For dichotomous multiple-choice and dichotomized constructed-response items, a three-parameter IRT model was used. Extended constructed-response and cluster items were polytomously scored and were analyzed with a generalized partial credit model (Muraki & Bock, 1991). The item parameter estimation was done separately within grade, but the final proficiency estimates were transformed to conform with the cross-grade scale created for the 1990 assessment. Within each grade, items were scaled using the full sample of students available from the 1990 and the 1992 assessments. The 1990 data were included at the scaling step of the analysis to assure that parameter estimates for items administered at both time points would not drift drastically between assessments. Item parameters were estimated separately for each of the five content area scales and the proficiency means and variances for subjects from the two assessment years were allowed to be different. In the final BILOG/PARSCALE run, the prior distributions of the population abilities were free to be estimated and the overall distribution was not centered at zero. In general, if an item was common to both assessment years, identical item parameters were used for both assessments. The appropriateness of the use of the identical parameters across assessments was examined by comparing the fit of the empirical item response functions against the estimated IRT item response functions. If IRT parameters did not fit the data, parameters specific to the assessment year were used. In the next step, sampling weights were applied and a final item estimation was run to convergence. (See Chapter 11 for further descriptions of the scaling process.) As with the long-term trend analysis, the calibration was based on student weights that were rescaled so that the their sum equaled the unweighted sample size of the 1992 sample. Also, weights for the 1990 data were restandardized to give equal weight to the two assessment years included in the scaling (see Appendix M).

Several items were deleted from the IRT analysis or were allowed to have different item parameters for different assessment years. Of the 368 total items in the main assessment, 12 (3 percent) received special treatment. These items are listed in Table 13-24, along with the reason for special treatment. If items had item response functions that were nonmonotonic, the items were deleted from scaling. This was the cases with only one item in the 1992 assessment. If item response functions for an item administered at different assessment years differed by year, the item was treated as if it were a different item in each assessment year and different item parameters were estimated. As it turned out, four items were treated in this way. Several items were responses to the same general question and had highly correlated responses. Such items were grouped together into a "cluster" item, which was scored polytomously, with the item score equal to the number of constituent cluster items the respondent answered correctly. Six items from the main assessment were cluster items.

For the special study involving estimation items, calibration was done separately by grade, although ability estimates were ultimately transformed to the 1990 cross-grade estimation scale. The entire sample for each grade was used. The sample weights were used so that estimates would accurately reflect the composition of the actual population. After evaluating the item response functions it was decided that three items, one from each grade, would be dropped for bad fit and that one item would have different item parameters for different assessment years (Table 13-25).

Table 13-25
IRT Scaling Results for the Estimation Scale:
Items Receiving Special Treatment

Grade	Item	Block	Subscale	Problem	Disposition
4	M032701	M16	Estimation	Different item responses for 1990 and 1992	1990 and 1992 response split into items M032701A and M032701B
	M032801	M16	Estimation	Was dropped in 1990	Dropped
8	M032101*	M16	Estimation	Was dropped in 1990	Dropped
12	M032101	M16	Estimation	Was dropped in 1990	Dropped

---

* All item parameters for grade 8 were estimated assuming a normal, 0,1 prior distribution for theta. The theta distribution for the unconstrained, nonnormal was very bimodal.

Appendix E lists the estimated item parameters that were produced by the BILOG/PARSCALE program.

Table 13-24
Items from the 1992 Mathematics Main Analysis Receiving Special Treatment

Grade	Item	Source	Block	Scale*	Problem	Disposition
4	M026531	1990 main	(1990 only)	A	Nonmonotonic item response function	Dropped
	M040401, M040402, M040403	1992 main	M9	B	Items highly correlated, not conditionally independent	Combined into cluster item M040461
	M044201N, M044202N	1992 main	M14	E	Items highly correlated, not conditionally independent	Combined into cluster item M044261N
	M044201O, M044202O	1992 calc. bridge	M14	E	Items highly correlated, not conditionally independent	Combined into cluster item M044261O
8	M028131	1990 main	(1990 only)	A	Nonmonotonic item response function	Dropped
	N275301	1990 main	(1990 only)	A	Nonmonotonic item response function	Dropped
	M022201	1990 main, 1992 main	M5	C	Different item responses for 1990 & 1992	1990 & 1992 responses split into items M022201A & M022201B
	M020901	1990 main, 1992 main	M6	C	Different item responses for 1990 & 1992	1990 & 1992 responses split into items M020901A & M020902B
	M050201 to M050204	1992 main	M3	D	Items highly correlated, not conditionally independent	Combined into cluster item M050261
	M045801 to M045804	1992 main	M7	D	Items highly correlated, not conditionally independent	Combined into cluster item M045861
12	M025301	1990 main, 1992 main	M5	A	Bad fit, dropped in 1990	Dropped
	M024901	1990 main, 1992 main	M5	B	Dropped in 1990	Dropped
	M020901	1990 main, 1992 main	M6	C	Different item responses for 1990 & 1992	1990 & 1992 responses split into items M020901A & M020901B
	M057401 to M057404	1992 main	M7	C	Items highly correlated, not conditionally independent	Combined into cluster item M057461
	M061101 to M061104	1992 main	M15	C	Items highly correlated, not conditionally independent	Combined into cluster item M061161
	M030731	1990 main	(1990 only)	D	Bad fit and very flat	Dropped
	M025401	1990 main, 1992 main	M5	E	Different item responses for 1990 & 1992	1990 & 1992 responses split into items M025401A & M025401B

---

* A = Numbers and Operations; B = Measurement; C = Geometry; D = Data Analysis Statistics and Probability; and E = Algebra and Functions.

BEST COPY AVAILABLE

13.2.8 Evaluating the Fit of the IRT Model

During the course of estimating an IRT model, individual items were evaluated to determine how well the item response model fit the data. This was done by visual inspection of plots comparing empirically based and theoretical item response functions. Specifically, for dichotomous items these plots consisted of nonmodel-based estimates of the expected proportion correct for each level of proficiency compared to the proportion correct for each level of proficiency as predicted by the theoretical item response function. For polytomous extended constructed-response items, similar plots were produced for each item category response function (see Chapter 9 for a fuller explanation of these plots).

In making decisions about excluding items from the final scales, a balance was sought between being too stringent, hence deleting too many items and possibly damaging the content representativeness of the pool of scaled items, and being too lenient, hence including items with model fit poor enough to endanger the types of model-based inferences made from NAEP results. Items that clearly did not fit the model were not included in the final scales; however, a certain degree of misfit was tolerated for a number of items included in the final scales.

For most items, the model fit well. In a few cases, poor fit with the data led to special treatment or deletion of the item. Figures 13-1 to 13-3 give examples of items with poor fit. In the item response plots the x-axis depicts proficiency, and the y-axis the probability of a correct response. The solid line is the logistic model prediction, and the symbols are the nonmodel-based predicted proportions. The size of the symbols are proportional to the estimate of the proficiency density. The overall proportion correct (p+) is given in the upper left corner, and the item parameter values are at the bottom right of the plot. For dichotomous items there were two sorts of poor fit: bad fit for one assessment and bad fit across two assessments. An example of bad fit for one assessment is found in Figure 13-1, which shows a nonmodel expected proportion correct that is nonincreasing, a definite violation of the item response model. This item was dropped. An example of bad fit across assessments is shown in Figure 13-2, in which the nonmodel item responses fall on either side of the theoretical line. This situation is remedied by relaxing the model to allow for a separate set of item parameters for each assessment year. The top and bottom halves of Figure 13-3 show the result of estimating separate parameters respectively for years 1990 and 1992. The nonmodel and theoretical item response functions closely agree in both cases.

For polytomously scored items there are three possible cases of bad fit: 1) Bad fit with a single assessment leading to the deletion of the item. There was no example of this with the mathematics assessment. 2) Bad fit with a single item which can be remedied by collapsing some of the response categories. Although the mathematics assessment had no such items, an example can be found in Chapter 12, section 12.2.4.2. 3) Bad fit across assessment years. There are no such examples since this is the first assessment employing polytomous IRT modeling.

13.2.9 Derived Background Variables

Derived variables based upon background questions were used for two purposes: as conditioning variables and as reporting variables used to define subgroups. Some of these

Figure 13-1

Example of Poor Item Fit: One Group

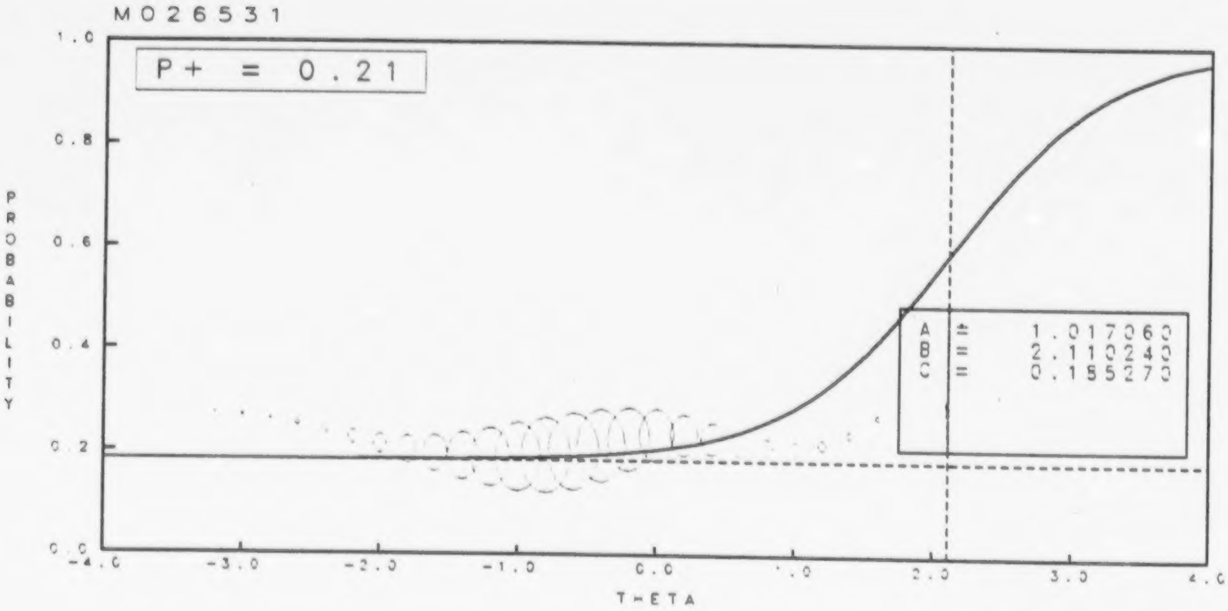
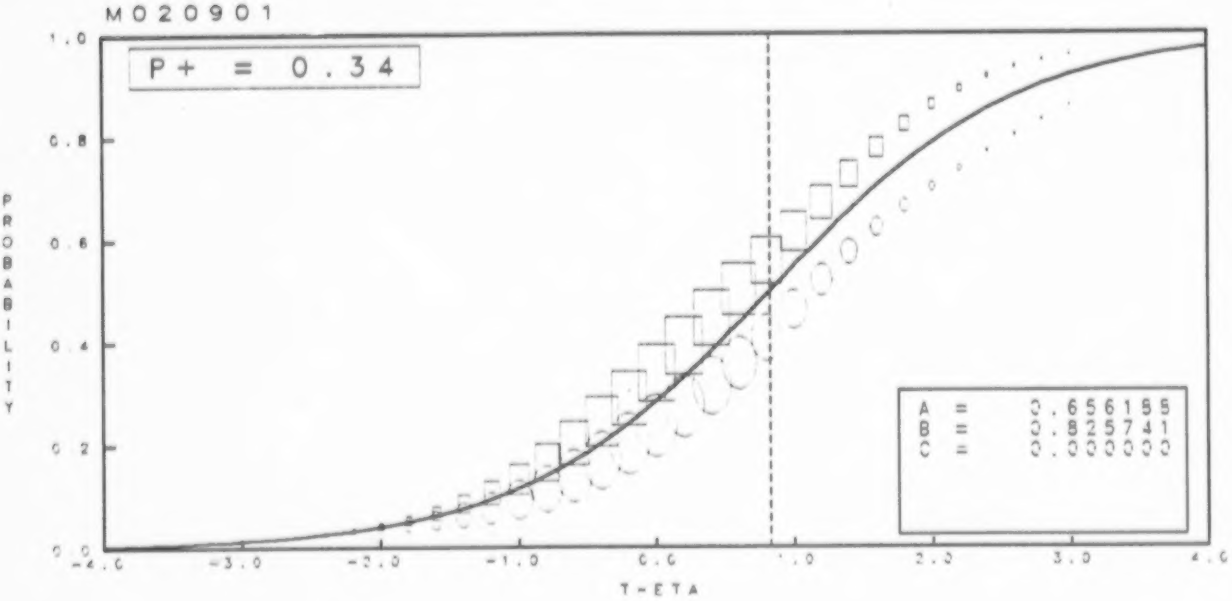


Figure 13-2

Example of Poor Item Fit: Two Subgroups*

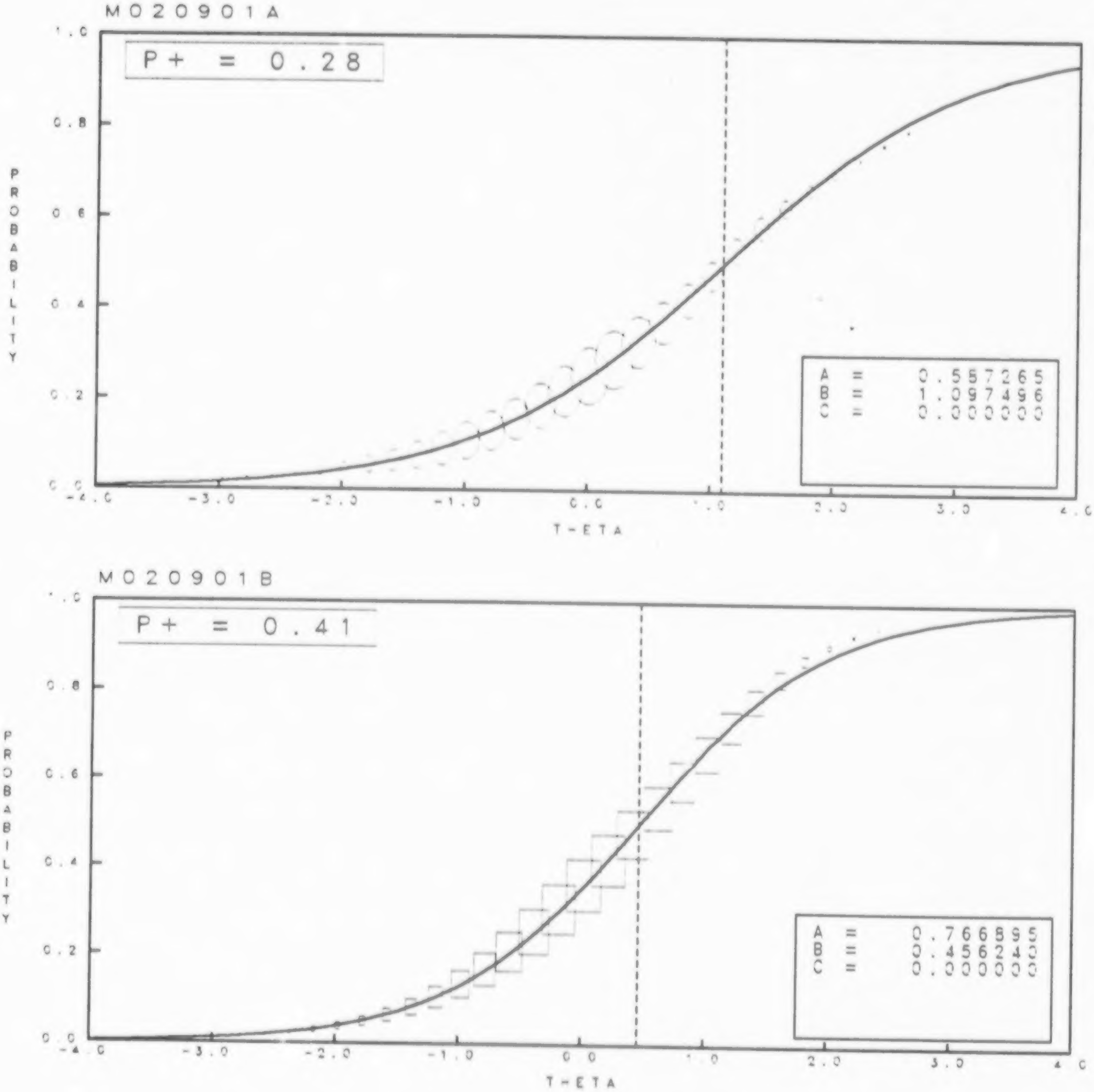


---

* Rectangles represent 1992 data; ovals represent 1990 data.

Figure 13-3

Examples of Fitting a Previously Poor-fitting Item to Two Separate Subgroups*



---

* Rectangles represent 1992 data; ovals represent 1990 data.

variables are common to all the subject areas; others are specific to the 1992 mathematics assessment. Derived variables used for conditioning and reporting are described in Appendix B.

13.2.10 Generation of Plausible Values

For the entire sample, multivariate plausible values for content area scales were generated for each age/grade group separately using the multivariate conditioning program CGROUP as revised by Thomas (1992). Final student weights were used at this stage of the analysis. Instead of using selected background variables for conditioning variables (as had been done prior to the 1990 assessment), principal components of the background variables were used. The principal components used accounted for 90 percent of the variance of the original conditioning variables. Principal components were employed to remedy problems of extreme collinearity among some of the original conditioning variables. For the estimation scale, univariate plausible values were generated for each separate age/grade; otherwise, the same procedures were followed for generating plausible values as for the main administration.

Research based on data from the 1990 Trial State Assessment in mathematics suggests that results obtained using such a subset of components will differ only slightly from whose obtained using the full set (Mazzeo, Johnson, Bowker, & Fong, 1992). Table 13-26 contains a list of the number of principal components included in conditioning, as well as the proportion of variance accounted for by the conditioning model for each age/grade.

Table 13-26
Proportion of Proficiency Variance Accounted for by the Conditioning Model
for the Mathematics Main Assessment

Grade	Number of Conditioning Contrasts	Number of Principal Components	Proportion of Proficiency Variance by Content Area Scale				
			Numbers and Operations	Measurement	Geometry	Data Analysis, Statistics, and Probability	Algebra and Functions
4	266	143	.42	.43	.51	.39	.42
8	311	160	.35	.37	.43	.37	.34
12	256	138	.31	.26	.29	.30	.20

The codings of the original mathematics-specific conditioning variables, before principal components were calculated, are presented in Appendix F. (For age 17/grade 12, the "modal age, > modal grade" category was deleted from the age-by-grade variable, because students above grade 12 were not sampled.) NAEP-CGROUP creates predictive distributions of proficiencies by combining information from item responses of individuals and information from linear regression of proficiency on conditioning variables. For each individual, five plausible values are randomly drawn from their predictive proficiency distribution.

The proportion of variance of each original conditioning variable accounted for by the principal components included in the conditioning model is listed in Appendix F. The estimated conditioning effects for the principal components of the three samples defined by the three age/grade groups are also given in Appendix F. The values of the conditioning effects are expressed in the metrics of the original calibration scale. Definitions of derived conditioning variables are given in Appendix B.

The CGROUP program (Mislevy, 1991; Rogers, 1991; Sheehan, 1985) which is used to condition the data, was revised (Thomas, 1992) to yield more accurate estimates of variances and covariances of the scales. In order to determine the effects of these revisions, the 1990 main assessment mathematics data were reanalyzed using the enhanced CGROUP program. Means were virtually unchanged, variances were moderately different, and covariances were significantly different from the original 1990 analyses. It was feared that differences in scale variances would significantly alter the transformation constants for the 1992 data. As a result, for the purposes of calculating coefficients for the 1992 data transformation, proficiencies were recalculated for the 1990 data using the revised CGROUP program.

For the 1990 assessment data that have been previously reported and that remain on secondary-use files, it was not deemed necessary to recalculate proficiencies using the enhanced CGROUP program because marginal means and variances for content area scales and composites were not greatly affected by the change of software.

13.2.11 The Transformation of the Proficiency Scale for Reporting and the Formation of the Composite Scale

Like all IRT scales, the mathematics content area scales have a linear indeterminacy that may be resolved by an arbitrary choice of the origin and unit-size in each given scale. In 1990 the NAEP mathematics data were scaled across grades separately for each scale. The linear indeterminacies among the scales were resolved by transforming the scale means and variances of three age/grade samples combined together to the 250.5, 50.0 metric using the transformation

$$\theta_{\text{proficiency}} = A \bullet \theta_{\text{calibrated}} + B.$$

As a result, all of the scales that spanned all three age/grade samples were on a common scale. By contrast, in 1992 each scale was scaled within grade. These scales were put on the 1990 cross-grade proficiency scale by a combination of common item and common population equating. The details of this linking are given below.

The IRT calibration was run with 1992 data and a comparable set of 1990 data from the winter assessment. Because there are items in common to the two assessments, this joint calibration put the 1992 data on the 1990 scale, but only up to a linear transformation. After the 1992 data was conditioned, the resulting individual thetas were also, by IRT assumptions, on the 1990 metric, except for a linear transformation.

The 1992 data was put on the 1990 metric by using a linear transformation that converts the 1992 thetas to the 1990 proficiency scale. This linear transformation was found by the following procedure. The 1990 thetas were reconditioned using CGROUP with the 1992 item parameters and the 1992 CGROUP conditioning parameters (gammas). This analysis resulted in 1990 thetas which were in the 1992 metric. We then transformed the new 1990 thetas (in the

1992 metric) to the 1990 proficiency metric (in the 1990 metric) by matching means and variances (i.e., linearly transforming the thetas so that they have the same mean and variance as the proficiencies). This is a common population equating procedure. The linear constants of this transformation were then used to transform the 1992 thetas to the 1990 proficiency metric. The transformation constants used for the five content area scales and for the estimation scale are given in Table 13-27. Note that comparable grade-only, winter-only subsample of the 1990 data was used for calculating the transformation.

Table 13-27
Coefficients of the Linear Transformations That Transform the Five Content Area Scales
from the 1992 Calibration Metric to the 1990 Reporting Proficiency Metric

Grade	Coefficient	Numbers and Operations	Measurement	Geometry	Data Analysis, Statistics, and Probability	Algebra and Functions
4	B	207.65	216.12	214.98	212.83	212.67
	A	37.59	35.87	30.81	34.59	32.10
8	B	264.59	258.17	258.18	261.10	259.67
	A	37.24	44.51	35.25	42.95	37.84
12	B	291.53	289.71	294.54	291.56	293.09
	A	34.48	40.10	40.22	37.16	39.20

The data analysis, statistics, and probability scale at grade 4 was not available for 1990. As a result, the 1992 grade 4 data analysis, statistics, and probability scale was transformed to have a mean and variance equal the average mean and variances of the other four transformed scales in the mathematics composite. This was plausible since the scales were highly correlated. Also, the data analysis, statistics, and probability scale had a distribution almost identical to the other scales in 1992.

The estimation scale was calibrated and transformed separately by a procedure similar to that for the other scales. The transformation constants for each grade are given in Table 13-28.

Table 13-28
Coefficients of the Linear Transformations That Transform the Estimation Scale
from the 1992 Calibration Metric to the 1990 Reporting Proficiency Metric

Grade	Coefficient	Estimation
4	B	199.05
	A	35.86
8	B	264.05
	A	31.45
12	B	289.93
	A	28.76

While multiple proficiency scales provide useful and very revealing information about the relative relationships among subpopulations, a single index to summarize overall performance is useful and communicative. For that reason, a mathematics composite was defined as a weighted average of the results across content area scales. The emphasis associated with each content area scale does not remain the same across grades. Therefore, the weights assigned to compute the average of the estimated content area scale proficiencies differ by grade. The assigned weights reflect the relative importance of content area scales for a particular age/grade as specified in the objectives for the mathematics assessment (NAEP, 1988). The definition of weights for the composite in each age/grade is given in Table 13-29. The mean and standard deviations of the composite proficiency scales for all three grades is given in Table 13-30. Note that this composite can be compared with the 1990 composite for reasons given in section 13.2.17.

Table 13-29
Defining Weights for the Mathematics Composite by Age/Grade

Scale	Age 9/Grade 4	Age 13/Grade 8	Age 17/Grade 12
Numbers and Operations	45	30	25
Measurement	20	15	15
Geometry	15	20	20
Data Analysis, Statistics, and Probability	10	15	15
Algebra and Functions	10	20	25

Table 13-30
Means and Standard Deviations on the Mathematics Composite Scale

	All Five Plausible Values	
Grade	Mean	S. D.
4	218.5	32.4
8	267.7	36.6
12	298.7	34.4

13.2.12 Analysis of Short-term Trend

Substantial gains in mathematics performance from 1990 to 1992 were seen at all grade levels and were broadly present across demographic groups. Because of the magnitude of these gains, numerous analyses were conducted to verify that the gains were real and not due to extraneous factors related to the conduct or analyses of the assessment. Similar gains were observed in the Trial State Assessment grade 8 mathematics results, which were based on local administration. Smaller gains were also observed in the long-term trend assessment, which involved a different instrument and time and mode of administration. Thus the gain was unlikely to be due to the administration of the mathematics assessment.

Table 13-31 gives an indication that the gain in scores is not due to the scaling. This table lists gains in the average percent correct metric for individual blocks for comparable subsamples of the assessment (i.e., winter-only, grade-only).

Table 13-31
Average Percent Correct for the 1992 Assessment
Versus a Comparable Subset of the 1990 Assessment*

Grade	1990		1992		1992 minus 1990 P+ Difference
	Block	**P+****	**Block**	**P+****	
4	D	.42	D	.43	.01
	F	.40	E	.43	.03
	E	.37	F	.41	.04
	H	.57	H	.57	.00
Simple average of 1990 to 1992 differences.					.02
8	D	.54	D	.58	.04
	F	.65	E	.68	.03
	E	.52	F	.56	.04
	H	.48	H	.49	.01
Simple average of 1990 to 1992 differences.					.03
12	D	.65	D	.68	.03
	F	.53	E	.56	.03
	E	.52	F	.56	.04
	H	.51	H	.53	.02
Simple average of 1990 to 1992 differences.					.03

---

* Winter-only, grade-only main assessment, weighted with respondents who were presented items but gave no response deleted.

** P+ = number correct/(number correct + number wrong)

To study the question of whether changes in demographics of the 1992 samples relative to the 1990 samples could explain the increase, a series of partitioning analyses were conducted. These analyses resulted in estimates of the amount of performance change attributable to each of three sources: population shifts (with no change in performance of subpopulations; changes in performance within subpopulations (with no shift in the relative frequency of subpopulations); and the interaction of shifts in population and performance.

Tables 13-32 and 13-33 display the results of applying this technique for a series of demographic variables. To remove any possible effect of scale construction, the analysis was conducted on the percent of items correct within a block. Table 13-32 displays the results of analyses conducted without using the sampling weights; Table 13-33 shows the results of the same analyses weighted. Both tables show that the bulk of the change is purely change in performance not associated with sample demographics. Since similar results are observed for both unweighted and weighted data, the change is not due to weighting. (It should be noted

Table 13-32
Results of Partitioning Analysis for Change Between 1992 and 1990
in Unweighted Percent of Items Correct in a Block, Averaged Across Blocks

Grade	Variable	Total Change	Change...		
			Due to Performance Changes	Due to Population Changes	Due to Interaction
4	Race/ethnicity	2.9	2.3	0.5	0.0
	Type of community	2.9	2.6	0.4	-0.1
	Parents' education	2.9	2.6	0.3	0.0
	Class ability	1.9	1.8	-0.2	0.2
8	Race/ethnicity	2.8	2.7	0.1	0.0
	Type of community	2.8	2.5	0.2	0.1
	Parents' education	2.8	2.4	0.3	0.0
	Class ability	3.2	4.0	-0.6	-0.2
12	Race/ethnicity	3.0	3.1	0.0	-0.1
	Type of community	3.0	3.2	-0.2	0.1
	Parents' education	3.1	2.6	0.4	0.1
	High-school program	4.5	3.9	0.6	0.1

Table 13-33
Results of Partitioning Analysis for Change Between 1992 and 1990
in Weighted Percent of Items Correct in a Block, Averaged Across Blocks

Grade	Variable	Total Change	Change...		
			Due to Performance Changes	Due to Population Changes	Due to Interaction
4	Race/ethnicity	2.8	2.7	0.1	0.0
	Type of community	2.8	2.6	0.2	0.0
	Parents' education	2.8	2.4	0.4	0.0
	Class ability	2.0	2.0	-0.2	0.1
8	Race/ethnicity	2.8	2.8	0.0	0.0
	Type of community	2.8	2.7	0.0	0.1
	Parents' education	2.8	2.5	0.3	0.0
	Class ability	3.5	4.0	-0.4	-0.2
12	Race/ethnicity	2.7	2.8	-0.1	0.0
	Type of community	2.7	2.4	0.2	0.1
	Parents' education	2.7	2.2	0.4	0.1
	High-school program	4.1	3.7	0.4	0.0

that the total change varies from variable to variable because not every student responded to every variable.)

Other analyses conducted are described in a letter by Design and Analysis Committee member Albert Beaton and the written response by Eugene Johnson and John Mazzeo. These appear in Appendix O.

In summary, after considering such factors as differential not-reached rates, absence, excluded students, and session monitoring (for the Trial State Assessment) none of the studied variables provided alternative explanations for the 1990-to-1992 gains.

13.2.13 Partitioning of the Estimation Error Variance

For each scale within each grade, the error variance of the transformed proficiency means was partitioned according to the procedure described in Chapter 11. The variance is partitioned into two parts; the proportion of error variance due to sampling students (sampling variance) and the proportion of error variance due to the fact that proficiency, θ, is a latent variable that is estimated rather than observed. Table 13-34 contains estimates of the total error variance, the proportion of error variance due to sampling students and the proportion of error variance due to the latent nature of θ (for stability of the estimates, these are based on 100 plausible values). More detailed information by gender and race/ethnicity is presented in Appendix N.

13.2.14 Mathematics Teacher Questionnaire

Teachers of fourth- and eighth-grade students assessed in mathematics were surveyed. Variables derived from the questionnaire were used in the conditioning models for the grade 4 and the grade 8 samples, along with a variable that indicated whether a student record had been matched with a teacher record, so that means for subgroups defined by these variables could be compared with no bias. Questionnaires were matched with students for 968 fourth-grade and 835 eighth-grade teachers. Of the 7,176 fourth-grade students in the main focused-BIB sample, 5,373 (75%) were matched with both parts of the teacher questionnaire and 1,036 (15%) were matched with only the first part of the questionnaire. Of the 7,663 eighth-grade students in the main focused-BIB sample, 6,105 (80%) were matched with both parts of the teacher questionnaire and 749 (10%) were matched with only the first part of the questionnaire. Thus, 89 percent of the fourth graders and 89 percent of the eighth graders were matched with at least the background information about their mathematics teachers.

13.2.15 Analysis of Dimensionality

As mentioned earlier, the main assessment is multivariate with five content area scales. Tables 13-35, 13-36, and 13-37 give conditional and marginal correlations for the five scales for the three grades. The conditional correlations are equivalent to correlations pooled within the demographic subgroups corresponding to grouping variables used to condition the data with CGROUP. Conditional correlations are the error correlations produced by a CGROUP

Table 13-34
Estimation Error Variance and Related Coefficients for the Mathematics Main Assessment

Grade	Scale	Total Estimation Error Variance	Proportion of Variance Due to...	
			Student Sampling	Latency of θ
4	Numbers and operations	0.68	0.87	0.13
	Measurement	0.75	0.83	0.17
	Geometry	0.49	0.78	0.22
	Data analysis, stat., & prob.	0.67	0.81	0.19
	Algebra and functions	0.75	0.76	0.24
	Composite	0.55	0.95	0.05
	Estimation	2.44	0.79	0.21
8	Numbers and operations	0.71	0.91	0.09
	Measurement	1.38	0.85	0.15
	Geometry	0.88	0.86	0.14
	Data analysis, stat., & prob.	1.15	0.85	0.15
	Algebra & functions	1.01	0.85	0.15
	Composite	0.80	0.98	0.02
	Estimation	1.58	0.89	0.11
12	Numbers and operations	0.70	0.85	0.15
	Measurement	0.93	0.84	0.16
	Geometry	1.01	0.86	0.14
	Data analysis, stat., & prob.	1.05	0.81	0.19
	Algebra and functions	0.99	0.94	0.06
	Composite	0.77	0.97	0.03
	Estimation	1.27	0.89	0.11

Table 13-35
Conditional and Marginal Proficiency Correlations, Grade 4

Scale	Numbers and Operations	Measurement	Geometry	Data Analysis, Statistics, and Probability	Algebra and Functions
Conditional Proficiency Correlations					
Numbers and Operations	1.00				
Measurement	.85	1.00			
Geometry	.73	.74	1.00		
Data Analysis, Stat., & Prob.	.92	.87	.77	1.00	
Algebra and Functions	.97	.87	.76	.92	1.00
Marginal Proficiency Correlations					
Numbers and Operations	1.00				
Measurement	.89	1.00			
Geometry	.82	.82	1.00		
Data Analysis, Stat., & Prob.	.94	.91	.85	1.00	
Algebra and Functions	.94	.88	.83	.92	1.00

Table 13-36

Conditional and Marginal Proficiency Correlations, Grade 8

Scale	Numbers and Operations	Measurement	Geometry	Data Analysis, Statistics, and Probability	Algebra and Functions
Conditional Proficiency Correlations					
Numbers and Operations	1.00				
Measurement	.87	1.00			
Geometry	.78	.82	1.00		
Data Analysis, Stat., & Prob.	.94	.84	.79	1.00	
Algebra and Functions	.95	.86	.85	.94	1.00
Marginal Proficiency Correlations					
Numbers and Operations	1.00				
Measurement	.91	1.00			
Geometry	.88	.89	1.00		
Data Analysis, Stat., & Prob.	.95	.90	.88	1.00	
Algebra and Functions	.95	.90	.90	.95	1.00

Table 13-37

Conditional and Marginal Proficiency Correlations, Grade 12

Scale	Numbers and Operations	Measurement	Geometry	Data Analysis, Statistics, and Probability	Algebra and Functions
Conditional Proficiency Correlations					
Numbers and Operations	1.00				
Measurement	.88	1.00			
Geometry	.80	.91	1.00		
Data Analysis, Stat., & Prob.	.91	.82	.77	1.00	
Algebra and Functions	.88	.87	.85	.89	1.00
Marginal Proficiency Correlations					
Numbers and Operations	1.00				
Measurement	.93	1.00			
Geometry	.90	.94	1.00		
Data Analysis, Stat., & Prob.	.93	.89	.88	1.00	
Algebra and Functions	.92	.91	.93	.91	1.00

analysis. We see that the conditional correlations are quite high, averaging .84 for grade 4, .86 for grade 8, and .86 for grade 12. The marginal correlations are the correlations of the first plausible values of each scale generated by CGROUP. Since background groupings are not controlled for, marginal correlations tend to be larger than conditional correlations, averaging .88 for grade 4, .91 for grade 8, and .91 for grade 12. Although it is of substantive interest to analyze the scales separately, the correlations indicate that they are highly redundant.

13.2.16 Anchoring the Points on the Mathematics Proficiency Scale

Scale anchoring was devised to associate descriptive statements of a student's ability with a level on a continuum of proficiency. This was done successfully with the 1986 reading, mathematics, and science scales. The same technique was applied to the 1992 mathematics composite scale. Four levels, 200, 250, 300, and 350, were selected on the scale and chosen as anchor points. Each level was defined by a description of the types of questions that most students attaining that level of proficiency would be able to answer correctly while most students at least one level lower would answer incorrectly. In this way each level was exemplified by typical benchmark items that describe a subset of abilities indicative of that level of proficiency.

The anchor points were defined in terms of the composite proficiency score. The anchoring was based on the empirical percent correct scores of items selected at each anchor point. The empirical percent correct was calculated by selecting subjects who had a proficiency score within 12.5 points of the anchor point score. After subjects were selected, their responses on items were averaged. For example, with regard to anchor point 300, students who scored between 287.5 and 312.5 on the proficiency scale (within 12.5 units of 300) were selected, if they were administered a particular item. These students were then used to estimate the average percent correct for items. To avoid problems of instability of estimates, if there were fewer than 100 students at a given proficiency range for an anchor point, the average percent correct statistic was not calculated. For further details on scale anchoring, see Beaton and Allen (1992).

As a result of the scale anchoring process for the mathematics composite, NAEP identified sets of items from the 1992 assessment that were good discriminators between subjects at different anchor proficiency levels. The guideline used to select such items was that students at any given anchor point would have a percent correct of at least 65 (but often higher) on the mathematics questions, while the students at the next lower level would have a much lower percent correct (50 or less) using the criterion that the difference in percents exceeds 30 between adjacent anchor points. In the case of the first anchor point, 200, the only criteria for selecting items was that subjects at that proficiency level have a minimum average percent correct of 65. Mathematics educators examined these sets of empirically selected items and used their expert judgment to characterize each proficiency level at each anchor point, contrasting tasks at that level with those at the levels just above and below. As a check on the generalizability of the interpretation process, the proficiency levels were independently described by two 10-member groups of mathematics educators. Upon comparing the results, both groups agreed that the two sets of interpretations were very similar and that either set would have appropriately described the anchor item information. The two groups then worked together to obtain the final interpretations.

The anchor levels were defined as follows:

200 - Addition and subtraction and simple problem solving with whole numbers.
250 - Multiplication and division, simple measurement, and two-step problem solving.
300 - Reasoning and problem solving involving fractions, decimals, percents, elementary concepts in geometry, statistics, and algebra.
350 - Reasoning and problem solving involving geometric relationships, algebra, and functions.

13.2.17 Grade 4: Comparing the Four-scale Composite of 1990 with the Five-scale Composite of 1992

The 1990 twelfth grade and eighth grade composites were composed of five content area scales, but the fourth grade composite involved only four scales, since it lacked the Data Analysis, Statistics, and Probability scale. In 1992, however, the Data Analysis, Statistics, and Probability scale exists for the fourth grade in both the main assessment and the Trial State Assessment. In order to assess the feasibility of using the five-scale composite of 1992, two composites were created for the 1992 main fourth-grade assessment:

1) a four-scale composite that is similar to the 1990 fourth grade composite, and

2) a five-scale composite that includes the Data Analysis, Statistics, and Probability scale. This composite is similar to the one used in the 1992 Trial State Assessment.

Means and jackknife standard error estimates were compared for these two composites and they were found to be virtually identical. As a result, the five-scale composite was used in all reporting of grade 4 results. Table 13-38 compares the distributions of the two composites for gender and race/ethnicity groups. Very little difference between the composites is evident.

Table 13-38

Grade 4: Comparing a Four-scale Composite with a Five-scale Composite at the Mean and Selected Percentiles

Subgroup Comparison	Mean	S.D.	Percentiles		
			5th	50th	95th
TOTAL					
Four-scale	218.4	32.5	162.3	220.0	269.4
Five-scale	218.5	32.4	162.4	220.1	269.4
Difference	-0.1	0.1	-0.1	-0.1	0.0
GENDER					
Male	219.7	33.4	161.9	221.5	271.9
Four-scale	219.7	33.3	162.0	221.6	271.8
Five-scale	0.0	0.1	-0.1	-0.1	0.1
Difference					
Female					
Four-scale	217.1	31.6	162.7	218.5	266.6
Five-scale	217.3	31.5	162.8	218.7	266.5
Difference	-0.2	0.1	-0.1	-0.2	0.1
RACE/ETHNICITY					
White					
Four-scale	226.7	29.1	176.5	227.9	272.6
Five-scale	226.8	29.1	176.7	228.1	272.5
Difference	-0.1	0.0	-0.2	-0.2	0.1
Black					
Four-scale	191.5	28.8	143.5	191.9	238.4
Five-scale	191.5	28.8	143.6	191.8	238.4
Difference	0.0	0.0	-0.1	0.1	0.0
Hispanic					
Four-scale	200.6	30.2	149.0	201.1	249.9
Five-scale	200.8	30.2	149.6	201.5	249.8
Difference	-0.2	0.0	-0.6	-0.4	0.1

Chapter 14

DATA ANALYSIS FOR THE SCIENCE LONG-TERM TREND ASSESSMENT[1]

Nancy L. Allen and Steven P. Isham

Educational Testing Service

This chapter describes the analyses performed on the responses to the cognitive and background items in the 1992 assessment of science. These analyses led to the results presented in *Trends in Academic Progress: Achievement of U.S. Students in Science, 1969 to 1992; Mathematics, 1973 to 1992; Reading, 1971 to 1992; and Writing, 1984 to 1992* (Mullis, Dossey, Campbell, Gentile, O'Sullivan, & Latham, 1994). The emphasis of this chapter is on the methods and results of procedures used to develop the IRT-based scale scores that formed the basis of these reports. The theoretic underpinnings of the IRT and plausible value methodology described in this chapter are given in Chapter 11.

The objective of the science analyses was to prepare scale values and perform all analyses necessary to produce a long-term trend report in science. The science trend line includes the years 1970, 1973, 1977, 1982, 1986, 1990, and 1992.

The student samples that were administered science items in the 1992 assessment are shown in Table 14-1. (See Chapters 1 and 3 for descriptions of the target populations and the sample design used for the assessment.)

Table 14-1
NAEP 1992 Science Student Samples

Sample	No. of Booklets	No. of Blocks	Booklets	Mode	Cohort Assessed	Time of Testing	Age Definition	Modal Grade	Number Assessed
9 [MS-LTTrend]	3	9*	91-93	Tape	Age 9	Winter	CY	4	7,335
13 [MS-LTTrend]	3	9*	91-93	Tape	Age 13	Fall	CY	8	5,909
17 [MS-LTTrend]	2	6*	84-85	Tape	Age 17	Spring	Not CY	11	4,359

LEGEND:

MS	Mathematics and science
LTTrend	Long-term trend assessment
Tape	Audiotape administration
CY	Calendar year: birthdates in 1982, 1978, and 1974 for ages 9, 13, and 17
Not CY	Age 17 only: birthdates between Oct. 1, 1974 and Sept. 30, 1974

* Of these, three are blocks of science items

---

[1] Data analysis and scaling were performed by Steven P. Isham.

The science trend results reported in *Trends in Academic Progress* are based on paced-tape administrations and occur at all of the age levels. The samples involved in the analysis were samples 9[MS-LTTrend], 13[MS-LTTrend], and 17[MS-LTTrend] in Table 14-1. For ages 9 and 13, the trend booklets for these samples contained blocks of reading, mathematics, and science items. The science and mathematics blocks were paced by tape-recordings and the reading blocks were presented in print form. The age 17 trend booklets contained only mathematics and science blocks, both presented by paced-tape recordings. All students received a block of common background questions, distinct for each age. Subject-area background questions were presented in the cognitive blocks. The booklets for the age 9 and age 13 samples (booklets 91-93) and the booklets for the age 17 samples (booklets 84-85) are the same as those used for long-term trend assessments in 1986 and 1990. The booklets and the blocks within those booklets are listed in Tables 4-13, 4-14, and 4-15 of Chapter 4. Additional information about all of the items in these blocks is in Tables 4-16, 4-17, and 4-18 of that chapter. This chapter includes specific information about the trend items that were scaled.

Table 14-2 clarifies the relationships between the 1992 trend samples and samples from previous years. For all ages, the paced-tape bridge to the 1986 samples allow direct comparisons with 1990 samples and with 1986 long-term trend samples. There was also a tape administration in 1988 at ages 9 and 13 that was comparable to the other years. However, a tape administration was not conducted at age 17 in 1988. Instead, a noncomparable paper-based assessment was conducted. Hence, 1988 is not included as a point in the long-term trend reporting. In 1986, the science trend items were scaled with common items from the 1977 and 1982 assessments. Because of the small number of items in common with those in the 1970 and 1973 assessments, data from those assessments were not scaled, but were linked to the trend line through mean proportion correct for common items. The 1990 trend assessments were linked to the 1970, 1973, 1977, and 1982 assessments through the 1986 assessment. The 1992 trend assessments were linked to previous assessments through the 1990 trend assessment. Information about previous assessment years, including 1970 and 1973, is available in Chapter 11 of *Expanding the New Design: The NAEP 1985-86 Technical Report* (Yamamoto, 1988) and Chapter 14 of *The NAEP 1990 Technical Report* (Allen, 1992).

The numbers of scaled items for each age are presented in Table 14-3. As was done with previous long-term trend analyses, each age was scaled separately and the trend scales are univariate. Derivation of subscales for specific content areas was not feasible given the limited number of items presented to students in the trend samples. The number of items scaled in 1992 that were common across assessment years is in Table 14-4.

The steps in the science long-term trend analysis are documented in the following sections. As is usual in NAEP analyses, the first step was to gather item and block information. Next, the trend items were calibrated, derived background variables were calculated, and plausible values were generated after conditioning on available background variables. Finally, the scales were placed on the final science trend proficiency scale used in previous trend assessments.

Table 14-2

NAEP Science Samples Contributing to 1992 Long-Term Trend Results, 1970-1992

Cohort	Year	Sample	Subjects	Time of Testing	Mode of Administration	Age Definition	Modal Grade
Age 9	1970	Main	SWC	Winter	Tape	CY	4
	1973	Main	MS	Winter	Tape	CY	4
	1977	Main	Sci	Winter	Tape	CY	4
	1982	Main	MSC	Winter	Tape	CY	4
	1986	LTTrend*	MS	Winter	Mixed	CY	4
	1990	LTTrend*	MS	Winter	Mixed	CY	4
	1992	LTTrend*	MS	Winter	Tape	CY	4
Age 13	1970	Main	SWC	Fall	Tape	CY	8
	1973	Main	MS	Fall	Tape	CY	8
	1977	Main	Sci	Fall	Tape	CY	8
	1982	Main	MSC	Fall	Tape	CY	8
	1986	LTTrend*	MS	Fall	Mixed	CY	8
	1990	LTTrend*	MS	Fall	Mixed	CY	8
	1992	LTTrend*	MS	Fall	Tape	CY	8
Age 17	1970	Main	SWC	Spring	Tape	Not CY	11
	1973	Main	MS	Spring	Tape	Not CY	11
	1977	Main	SL	Spring	Tape	Not CY	11
	1982	Main	MSC	Spring	Tape	Not CY	11
	1986	LTTrend*	MS	Spring	Tape	Not CY	11
	1990	LTTrend*	MS	Spring	Tape	Not CY	11
	1992	LTTrend*	MS	Spring	Tape	Not CY	11

LEGEND:

Sci	Science
MS	Mathematics and science
MSC	Mathematics, science, and citizenship
SL	Science, life skills
SWC	Science, writing, and citizenship
Main	Main assessment

* Within a cohort, these samples received common booklets

LTTrend	Long-term trend assessment
Tape	Audiotape administration
Mixed	Mathematics and science administered by audiotape, reading administered by print
CY	Calendar year: birthdates in 1982, 1978, and 1974 for ages 9, 13, and 17
Not CY	Age 17 only: birthdates between Oct. 1, 1974 and Sept. 30, 1974

Table 14-3

Numbers of Scaled Science Long-term Trend Items Common Across Ages

Age	Booklets	Number of Items
9 only	91-93	55
13 only	91-93	30
17 only	84-85	32
9 and 13 only	91-93, 91-93	0
9 and 17 only	91-93, 84-85	0
13 and 17 only	91-93, 84-85	45*
9, 13, and 17	91-93, 91-93, 84-85	1
Total		163

---

* One of these items (N406303) was treated as a different item in the scaling of the 1992 assessment, but only for the age 13 students.

Table 14-4

Numbers of Scaled Science Long-term Trend Items Common Across Assessments

Assessment Years	Number of Items		
	Age 9	Age 13	Age 17
1986, 1990, 1992	56	76	78
1982, 1986, 1990, 1992	10*	58	47
1977, 1986, 1990, 1992	56	76	76
1977, 1982, 1986, 1990, 1992	10*	58**	45

---

* Twenty-four items common to years 1977 and 1982, but not later years, were included in the 1986 scaling of these items to stabilize the estimation of the item parameters. See *Expanding the New Design: The NAEP 1985-86 Technical Report* for more information.

** One of these items (N406303) was treated as a different item when scaled for the 1992 assessment.

14.1 Item Analysis for the Long-term Trend Assessment

Conventional item analyses did not identify any difficulties with the 1992 long-term trend data for the 1992 samples that bridge to 1986. Table 14-5 contains information about the science trend blocks. These blocks were presented to samples 9[MS-LTTrend], 13[MS-LTTrend], and 17[MS-LTTrend]. At ages 9 and 13, the blocks labeled S1, S2, and S3 were presented intact to 1986 and 1990 trend samples. The age 9 and age 13 blocks appeared in booklets 91 through 93. At age 17, S1, S2, and S3 were presented intact to the 1986 and 1990 trend samples. Block S3 was in booklet 84 and blocks S1 and S2 were in booklet 85. Table 14-6 shows the relationships between the blocks and booklets. Common labeling of these blocks across ages does not denote common items.

Table 14-5 contains the number of items, size of the sample administered the block, mean weighted proportion correct, mean weighted r-biserial, and mean weighted alpha as a measure of reliability for each block. The average values were calculated using examinee sampling weights and the items in the block that were scaled. On average, the 1992 item-level statistics were slightly higher, but not very different from those for the 1986 and 1990 assessments. There were a number of individual items that had larger weighted proportion-correct values for the 1992 sample than for 1990 sample of students. One of these, item N406303, was administered to age 13 students in block S1. This item is discussed in section 14.2. Larger weighted proportion-correct values for the 1992 sample for other items did not significantly affect the estimation of item parameters using data from the 1990 and 1992 assessments. The percent of examinees not reaching items in the bridge blocks was always zero because the items were administered with a tape-recording to pace response time.

14.2 Estimation of Item Parameters

The first step in the scaling process (described in Chapter 11) was the estimation of item parameters for the trend items. This item calibration was performed using the NAEP version (Rogers & Nelson, 1990) of the BILOG and PARSCALE programs (Mislevy & Bock, 1982; Muraki & Bock, 1993) separately for each of the three age groups, using combined data from the 1990 and 1992 assessment years and treating each assessment sample as a sample from a separate subpopulation. The calibration was performed on all examinees using student weights during the entire scaling process. The weights for the 1990 samples were used in a rescaled form, where the sum of the rescaled weights for the 1990 samples was equal to the sum of the weights for the 1992 samples.

Item parameters were estimated separately for items at each age using 1990 and 1992 data with the NAEP-BILOG/PARSCALE computer program. Although other items were examined for irregularities, only items that were deleted from the previous scaling of the paced-tape trend data were excluded in the 1992 analysis. Eleven percent of the items (18 items) administered to the long-term trend sample were excluded from analyses of previous assessments. The deleted items appear in Table 14-7. As a result of these deletions, 56 items were scaled for age 9, 76 items were scaled for age 13, and 78 items were scaled for age 17. A list of the items scaled for each of the ages, along with their item parameter estimates, appears in Appendix E.

Table 14-5

Descriptive Statistics for Item Blocks
Long-term Science Trend Samples

	Block		
Statistic	**S1**	**S2**	**S3**
Age 9			
Number of scaled items	17	20	19
Number of scaled constructed-response items	0	0	0
Unweighted sample size	2388	2512	2435
Average weighted proportion correct	.62	.58	.71
Average weighted r-biserial	.58	.48	.55
Weighted alpha reliability	.71	.64	.68
Age 13			
Number of scaled items	23	30	23
Number of scaled constructed-response items	0	0	0
Unweighted sample size	1928	1976	2005
Average weighted proportion correct	.54	.57	.61
Average weighted r-biserial	.51	.48	.53
Weighted alpha reliability	.72	.77	.72
Age 17			
Number of scaled items	24	31	23
Number of scaled constructed-response items	0	0	0
Unweighted sample size	2152	2152	2207
Average weighted proportion correct	.66	.65	.60
Average weighted r-biserial	.50	.53	.60
Weighted alpha reliability	.70	.78	.80

Table 14-6

Correspondence Between Samples, Booklets, and Blocks for the Science Long-term Trend

Sample	Booklet	Block
Age 9 MS-LTTrend	91	S1
	92	S2
	93	S3
Age 13 MS-LTTrend	91	S1
	92	S2
	93	S3
Age 17 MS-LTTrend	84	S3
	85	S1 S2

Table 14-7

Items Deleted from the Science Paced-tape Long-Term Trend Analysis

Age	Booklet	Block	Item	Reason for Exclusion
9	91	S1	N400201	Excluded in previous assessment
	92	S2	N401701	Excluded in previous assessment
	92	S2	N402003	Excluded in previous assessment
	92	S2	N402004	Excluded in previous assessment
	92	S2	N402601	Excluded in previous assessment
	92	S2	N402603	Excluded in previous assessment
	93	S3	N403802	Excluded in previous assessment
13	91	S1	N404902	Excluded in previous assessment
	91	S1	N404903	Excluded in previous assessment
	92	S2	N407501	Excluded in previous assessment
	93	S3	N409401	Excluded in previous assessment
	93	S3	N409402	Excluded in previous assessment
	93	S3	N409403	Excluded in previous assessment
	93	S3	N409801	Excluded in previous assessment
17	85	S1	N410001	Excluded in previous assessment
	85	S1	N410002	Excluded in previous assessment
	85	S1	N410301	Excluded in previous assessment
	85	S2	N407402	Excluded in previous assessment

BEST COPY AVAILABLE

The item response function, as well as the weighted proportion correct (as mentioned in section 14.1), for one item (N406303) presented to age 13 students in block S2 was different across the two assessment years included in the scaling. Figure 14-1 contains the empirical and model-based item response functions for this item. The solid curve represents the model-based item response function, while the diamonds represent the empirical item response function for 1990 and the ovals represent the item response function for 1992. A description of functions of this type are given in Chapter 9. Because the empirical item response functions for the two assessment years were so different, this item was treated as if it were a different item in each of the two years. It was not considered to be an item that was common to both the 1990 and 1992 assessments, as all of the other items were.

14.3 Derived Background Variables

In the long-term trend analysis, all variables derived for the science analysis from background questions were used both in generating plausible values and in reporting (to define subgroups). Derived conditioning and reporting variables are described in Appendix B.

14.4 Generation of Plausible Values

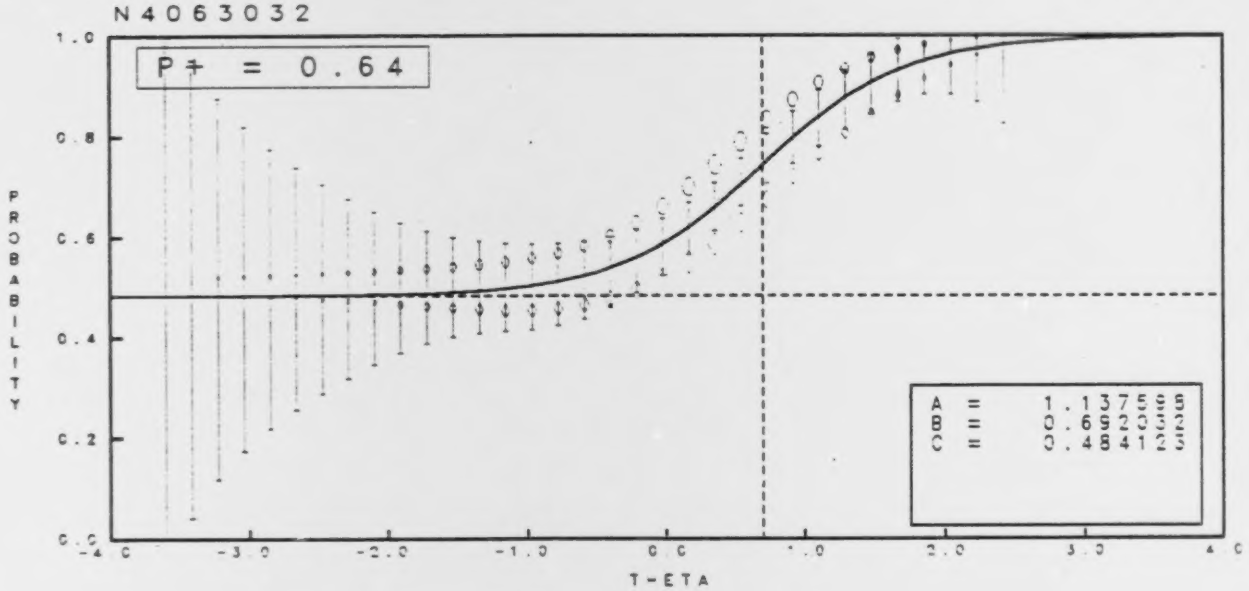
The generation of plausible values was conducted independently by age for each of the three assessment years. The item parameters from NAEP-BILOG/PARSCALE, final student weights, item responses and selected background variables (conditioning variables) were used with the computer program BGROUP (described in Chapter 11) in order to generate the values for each age. There were 45 contrasts in the conditioning model at age 9, including an overall constant, 48 at age 13, and 54 at age 17. Appendix F gives the codings for the conditioning variables and the estimated conditioning effects for the three age groups. The estimated conditioning effects in the tables are expressed on the scale of the original calibration. A check on the distributions of the plausible values for each age was made. The generation of plausible values is described in more detail in Chapters 9 and 11. Table 14-8 shows the proportion of variance accounted for by the conditioning model for each age.

Table 14-8
Proportion of Proficiency Variance Accounted for by the Conditioning Model
for the Science Long-term Trend Assessment

Age	Number of Conditioning Contrasts	Proportion of Proficiency Variance
9	45	.669
13	48	.572
17	54	.472

Figure 14-1

Plot Comparing Empirical and Model-based Estimates of the Item Response Function for Item N406303 in the 1992 Long-term Science Trend Samples*



---

* Solid curve represents model-based item response function; diamonds represent 1990 data; ovals represent 1992 data.

14.5 The Final Proficiency Scale

The linear indeterminacy of the trend scale was resolved by linking the 1990 trend scales to the previous long-term trend scales using the following procedure. For each age, the item parameters from 1992 based on data from 1990 and 1992 were used with the 1990 data to find plausible values for the 1990 data. The mean and standard deviation of all of the plausible values were calculated and matched to the mean and standard deviation of all of the plausible values based on the 1990 item parameters and 1990 data as reported in earlier reports. The transformations that resulted from this matching of the first two moments for the 1990 data are

Age 9: $\theta_{\text{proficiency}} = 33.26 \bullet \theta_{\text{calibrated}} + 232.60$

Age 13: $\theta_{\text{proficiency}} = 39.78 \bullet \theta_{\text{calibrated}} + 255.52$

Age 17: $\theta_{\text{proficiency}} = 45.04 \bullet \theta_{\text{calibrated}} + 292.39$

where $\theta_{\text{proficiency}}$ denotes values on the final transformed scale and $\theta_{\text{calibrated}}$ denotes values on the original calibration scale. Overall summary statistics for the long-term trend samples are given in Table 14-9.

Table 14-9
Means and Standard Deviations on the Science Long-Term Trend Proficiency Scale

		All Five Plausible Values	
Age	**Assessment**	**Mean**	**S. D.**
9	1990 1992	228.7 230.6	40.2 39.9
13	1990 1992	255.2 258.0	37.6 36.9
17	1990 1992	290.4 294.1	46.2 44.7

14.6 Partitioning of the Estimation Error Variance

The variance of proficiency means for each grade was partitioned into the part due to the sampling of students and the part due to the latency of proficiency, θ, as described in Chapter 11. These estimates are given in Table 14-10 (for stability of the estimates, they are based on 100 plausible values). More detailed information for gender and race/ethnicity subgroups is available in Appendix N.

Table 14-10
Estimation Error Variance and Related Coefficients for the Science Long-term Trend Assessment

Age	Total Estimation Error Variance	Proportion of Variance Due to...	
		Student Sampling	Latency of θ
9	0.70	0.81	0.19
13	0.71	0.87	0.13
17	1.66	0.92	0.08

14.7 Anchoring the Points on the Science Proficiency Scale

The main NAEP science composite scale was anchored in 1986, using the process described in *Expanding the New Design: The 1985-86 Technical Report*. Because each of the 1992 scales was tied to the 1986 main cross-sectional or long-term trend scale through the 1990 data, the distribution of proficiency scores derived from the main and bridge samples can be described in terms of scale anchors. In 1986 the levels of science proficiency were

150 - Knows everyday science facts;
200 - Understands simple scientific principles;
250 - Applies basic scientific information;
300 - Analyzes scientific procedures and data; and
350 - Integrates specialized scientific information.

Chapter 15

DATA ANALYSIS FOR THE WRITING ASSESSMENT[1]

James E. Carlson, Eiji Muraki, Bruce Kaplan, and Yim Fai Fong

Educational Testing Service

This chapter describes analyses of the writing prompts[2] and background items in the 1992 assessments of writing. These analyses led to the results reported in *NAEP 1992 Trends in Academic Progress* (Mullis, Dossey, Campbell, Gentile, O'Sullivan, & Latham, 1994), the *NAEP 1992 Writing Report Card* (Applebee, Langer, Mullis, Latham, & Gentile, 1994), and *Windows into the Classroom: NAEP's 1992 Writing Portfolio Study* (Gentile & Martin-Rehrmann, 1994). Emphasis is given to the psychometric methods used to develop the composite scores that formed the basis of those reports.

The objectives of the 1992 writing analyses were to:

- prepare scale values and perform analyses required to produce a long-term trend report in writing. The writing trend study currently includes the years 1984, 1988, 1990 and 1992.
- prepare scale values and perform analyses required to produce a report for the main assessment in writing.
- perform the analyses required to produce a report on the type of school-based writing that fourth- and eighth-grade students are doing as part of English/language arts instruction, and the characteristics and quality of that writing.

Each of these major analyses is discussed in a separate section of this chapter. Section 15.1 covers the trend analyses, section 15.2 the main assessment analyses, and section 15.3 the school-based writing analyses. Some aspects of the analyses that are the same as those used in other subject areas are not described here because they are discussed in Chapter 9 or a later chapter.

The specific 1992 samples used for the analysis of writing achievement by age cohort are presented in Table 15-1. The first three samples are those for the three cohorts of the main assessment. The other three samples are those selected for the writing trend assessment. These samples comprise students selected both on the basis of age and grade in school. For the writing trend, unlike other subject-area trend assessments, only those students selected on the basis of grade were included and the sample is referred to as a "grade-only" sample.

---

[1] Angela Grima also contributed to this chapter. Data analysis and additional statistical programming were performed by Lucie Chan, Phillip Leung, Michael Narcowich, and Ira Sample.

[2] The terms "item" and "prompt" are used interchangeably in this chapter.

Table 15-1
NAEP 1992 Writing Student Samples

Sample	# of Booklets	# of Blocks	Mode	Cohort Assessed	Time of Testing	Age Defn.	Modal Grade	Sample Size
9[Wrt-MainP]	18	9	Print	Age 9/grade 4	1/6/92 - 3/31/92	CY	4	9,552
13[Wrt-MainP]	20	9+2*	Print	Age 13/grade 8	1/6/92 - 3/31/92	CY	8	14,492
17[Wrt-MainP]	21	9+3*	Print	Age 17/grade 12	1/6/92 - 3/31/92	CY	12	15,669
9[RW-LTTrend]	6	13	Print	Age 9/grade 4	1/6/92 - 3/31/92	CY	4	7,062
13[RW-LTTrend]	6	14	Print	Age 13/grade 8	10/7/91 - 12/13/91	CY	8	5,514
17[RW-LTTrend]	6	14	Print	Age 17/grade 11	3/16/92 - 5/15/92	Not CY	11	5,569

LEGEND:

Wrt	Writing
RW	Reading and writing
MainP	Main assessment, print administration
LTTrend	Long-term trend assessment

CY	Calendar year: birthdates in 1982, 1978, and 1974 for ages 9, 13, and 17
Not CY	Age 17 only: birthdates between Oct. 1, 1974 and Sept. 30, 1975

* 50-minute blocks

A major departure from analysis procedures used in previous writing assessments was the use of IRT scaling for the first time. This necessitated the development of procedures that differ from those described for the writing analysis in previous technical reports. Sampling weights, as described in Chapter 10, were used for all analyses of the writing prompts.

15.1 LONG-TERM TREND DATA ANALYSIS

Because IRT scaling was used for the first time in 1992 for the writing long-term trend assessment, a new scale was developed and data from the 1984, 1988, 1990 long-term trend samples were reanalyzed in order to establish this scale. The 1992 writing trend points are based on data from the 1992 reading/writing long-term trend samples (RW-LTTrend in Table 15-1). The booklets used in this assessment contained blocks of reading and writing items, as well as background questions. Each subsequent assessment matched the 1984 assessment in terms of the time of administration and age definitions. Identical booklets were used in 1984, 1988, 1990, and 1992.

The items on which the trends in writing achievement are based are shown in Table 15-2. The table shows the block that contained the item in 1984 and trend booklets containing the item in 1988, 1990, and 1992. Twelve writing tasks were used to measure trends, with six tasks presented at each grade level. To allow comparisons in writing ability across grades, three of the six tasks presented to fourth-grade students were also presented to eighth-grade students;

three of the eighth-grade tasks were also presented to eleventh-grade students; and one of the common tasks (Appleby House) was presented at all three grade levels.

Table 15-2
Assignment of 1984-1992 Writing Long-term Trend Items in 1984, 1988, 1990, and 1992

Writing Task	1984 BIB-spiral Blocks Used for Trend Analyses			1988 to 1992 Trend Booklets		
	Grade 4	Grade 8	Grade 11	Grade 4	Grade 8	Grade 11
N0003 Recreation Opportunity		C	C		52,54	52,54
N0004 Food On Frontier		D	D		51,54	51,54
N0005 Dissecting Frogs		E			53,55	
N0006 XYZ Company	E	E		52, 54	53,55	
N0009 Radio Station	G	G		54,55	55,56	
N0010 Appleby House	G	G	G	54,55	55,56	55,56
N0076 Flashlight	V*			56		
N0147 Plants	C			51,53		
N0148 Spaceship	E			52,54		
N0180 Space Program			E			53,55
N0190 Job Application			E			53,55
N0210 Bike Lane			G			55,56

---

* Block V was not placed in a booklet with any other writing block in 1984 (all other blocks appeared with every other block at the same grade in 1984) and hence it could not be used in scaling.

15.1.1 Primary Trait Scoring of the Writing Tasks and Measures of Scorer Reliability

All writing exercises from the 1992 assessment were scored for task accomplishment (primary trait). For the purposes of analysis, the student responses were coded as 0 (not rated), 1 (unsatisfactory), 2 (minimal), 3 (adequate), and 4 (elaborated). "Not-reached" and "Omitted" items were excluded from the scaling. The writing trend blocks contained either one or two items. If an item was left blank in a one-item block, it was scored as an omission. Items considered not-reached occurred only in writing blocks that had two cognitive items in which the first item was answered and the second was not.

A 25 percent random subsample of all 1992 papers was rescored by a second rater to provide an estimate of interrater reliability. Although the measures of scorer agreement in NAEP have been consistently high, we recognized the possibility that there might be variation between the ratings provided by the group of scorers assembled in 1992 and the scorers assembled in previous years. Such a variation would be a confounding effect in the measurement of trend. The most direct way of controlling the effect of across-year variation in scoring would be to eliminate it entirely by rescoring all of the data from the previous three assessments, using the same set of scorers who scored the 1992 data. Unfortunately, resources did not allow for the rescoring of the full set of writing papers but did allow for a rescoring of over 11,000 of the papers from 1988 (the numbers by prompt and grade are displayed in Table 15-3). The rescored papers for a given item constituted approximately a 25 percent sample of all 1988 papers and consisted of all grade-eligible respondents to two or three of the 1988 booklets containing that item. This procedure of rescoring data from previous writing assessments was also used in the previous years' assessments used in the current trend analysis (1984, 1988, 1990).

It was expected that, because of rigorous training of scorers, the between-year variability in scoring would be low enough to permit the use of the full set of the 1988 data. Table 15-3 shows scorer reliability, as measured by the intraclass correlation, for each prompt in the 1988, 1990, and 1992 data. The percentage of exact agreement between first and second raters is also given. In addition, the table shows the intraclass correlation and percentage of exact score agreement comparing the scores of samples of the 1990 and 1992 raters with those of the 1988 raters on a sample of the 1988 papers. The reliabilities and percentages of exact agreement (between first and second raters) were generally high for 1992 data, as they were for prior assessments (1984, 1988, and 1990).

15.1.2 Item Analysis of the Writing Trend Items

A standard item analysis for polytomous items, as described in Chapter 9, was conducted on the writing trend item data. Table 15-4 displays the item analysis statistics for each grade—the number of examinees responding to each prompt, the percentage of examinees receiving each of the assigned scores, and the mean score of the prompt.

The results of the item analysis were examined to verify that statistics for each item were in expected ranges. No difficulties were found in this process. Comparisons of item statistics with those of previous years were also made, and it was found that the items had similar statistics for the four years of the long-term trend in writing.

15.1.3 Calibration of Writing Trend Prompts Using the Generalized Partial Credit Model

This section provides information regarding the scaling of the primary trait data from the 1992 writing trend assessment.

A listing of the prompts used in scaling at the three grade levels is presented in Table 15-5. Five prompts were used at the fourth grade and six prompts were used at each of the eighth- and eleventh grades. Either three or four scoring categories were used in the scaling for each of the prompts. Two prompts at the fourth grade and one at the eighth were scaled with three categories because the frequencies of responses in the fourth category were zero or near zero. All other prompts were scaled with four categories. After examining the pattern of omitted, not-reached, off-task, and illegible responses, relative to responses to other prompts, it

Table 15-3

Percentages of Exact Score Agreement and Interrater Reliability for the Primary Trait Scoring of the Writing Long-term Trend Assessment Items

NAEP Item	1988 Data (by 1988 Raters)		1988 Data (by 1990 Raters)		1990 Data (by 1990 Raters)		1988 Data (by 1992 Raters)		1992 Data (by 1992 Raters)	
	Percent Agreement	Reliability	Percent Agreement	Reliability	Percent Agreement	Reliability	Percent Agreement	Reliability	Percent Agreement	Reliability
Grade 4/Age 9										
N0006 XYZ Company	97.1	.99	91.1	.90	88.8	.83	90.9	.92	88.0	.89
N0009 Radio Station	93.5	.95	89.0	.90	92.1	.93	92.0	.93	90.9	.93
N0010 Appleby House	90.3	.92	76.9	.78	78.5	.72	79.7	.83	78.4	.80
N0076 Flashlight	87.5	.88	80.5	.74	78.2	.77	76.6	.71	79.5	.75
N0147 Plants	94.3	.95	88.5	.89	82.4	.86	91.3	.92	88.5	.90
N0148 Spaceship	91.8	.95	83.7	.89	75.2	.82	77.8	.84	75.7	.79
Grade 8/Age 13										
N0003 Recreation Opportunity	85.4	.82	83.0	.81	76.7	.73	79.5	.77	81.9	.82
N0004 Food on Frontier	79.9	.68	83.5	.78	72.1	.67	79.4	.68	75.5	.69
N0005 Dissecting Frog	76.1	.64	80.6	.70	66.1	.56	71.2	.54	71.7	.63
N0006 XYZ Company	93.5	.92	92.6	.87	86.8	.76	86.2	.76	88.5	.80
N0009 Radio Station	87.0	.89	82.0	.79	80.7	.83	85.8	.87	86.0	.89
N0010 Appleby House	75.3	.69	75.4	.75	75.9	.72	78.0	.77	82.5	.84
Grade 11/Age 17										
N0003 Recreation Opportunity	90.8	.93	71.6	.78	76.3	.78	83.7	0.84	79.9	.78
N0004 Food on Frontier	93.1	.86	78.9	.69	76.7	.73	85.9	0.72	79.7	.73
N0010 Appleby House	89.3	.89	81.1	.81	81.6	.82	88.2	0.88	81.4	.82
N0180 Spaceship	89.9	.93	73.2	.75	71.8	.75	82.9	0.86	80.8	.81
N0190 Job Application	92.3	.92	85.5	.86	84.6	.83	90.2	0.88	90.5	.84
N0210 Bike Lane	84.9	.87	78.2	.76	75.6	.78	83.5	0.84	81.5	.83

Table 15-4

Descriptive Statistics for Writing Prompts, Writing Long-term Trend Samples

Item	N	Percentages of Examinees in Each Score Category					Mean
		0	1	2	3	4	
Grade 4							
N0006	1583	2.2	50.7	11.6	35.4		1.80
N0009	1650	3.9	45.6	32.8	17.4	0.3	1.65
N0010	1337	1.7	22.8	53.0	22.4	0.1	1.97
N0076	850	0.8	28.7	58.3	11.9	0.3	1.82
N0147	1677	1.0	16.1	42.7	40.2		2.22
N0148	1653	4.7	31.4	43.9	19.6	0.4	1.80
Grade 8							
N0003	1317	1.2	50.4	33.8	13.8	0.8	1.63
N0004	1316	0.4	21.1	64.8	13.1	0.6	1.92
N0005	1359	0.9	11.6	67.2	19.1	1.2	2.08
N0006	1333	0.2	21.4	6.1	72.3		2.51
N0009	1362	0.8	27.5	40.2	30.0	1.4	2.04
N0010	1249	0.3	7.9	33.1	57.3	1.3	2.51
Grade 11							
N0003	1416	1.3	31.6	52.4	14.5	0.2	1.81
N0004	1447	1.2	13.9	68.0	16.1	0.7	2.01
N0010	1264	1.0	9.5	36.4	50.9	2.1	2.44
N0180	1427	3.1	14.3	55.7	25.6	1.3	2.08
W0190	1403	0.3	16.9	11.4	70.9	0.6	2.55
W0120	1425	1.7	33.7	43.1	19.8	1.7	1.86

Table 15-5

Sample Sizes for Scaling Items Used in Trends in Writing Performance

NAEP Item	Grade 4				Grade 8				Grade 11			
	1984 Rescore	1988	1990	1992	1984 Rescore	1988	1990	1992	1984 Rescore	1988	1990	1992
N0006 XYZ Company	603	1398	1679	2023	751	1766	1992	1787				
N0009 Radio Station	582	1506	1794	2099	713	1801	2025	1834				
N0010 Appleby House	530	1143	1413	1684	674	1658	1848	1681	694	1322	1628	1614
N0147 Plants	709	1604	1833	2213								
N0148 Spaceship	664	1508	1765	2076								
N0003 Recreation Opportunity					598	1758	1962	1729	635	1482	1767	1790
N0004 Food on Frontier					731	1748	1951	1735	714	1482	1756	1806
N0005 Dissecting Frogs					777	1796	2016	1816				
N0180 Space Program									713	1464	1776	1777
N0190 Job Application									728	1473	1796	1804
N0210 Bike Lane									748	1495	1823	1801

was decided to treat these responses as missing, because there appeared to be no obvious relationship between writing performance and nonresponse (for whatever reason) to other prompts. Treating such prompts as missing meant that they would not enter into the scaling. Note that one of the prompts administered in the assessment, "Flashlight," was not administered with any other prompt and therefore could not be put on the same scale as the other prompts. This prompt was therefore excluded from scaling, and from the number of prompts cited above.

The resultant sample sizes for scaling the trend items are provided in Table 15-5. The 1984 trend point was based on a rescoring of a sample of the 1984 data by the 1988 raters. For details, see Johnson (1990).

When the trend items were administered for the first time in 1984, they were used as part of the 1984 BIB design. By applying the additional information that could be obtained from the 1984 design, the trend items for each grade were calibrated together.

This task was not straightforward. Due to rater drift, the 1984 original data for trend items had to be rescored in 1988 and not all the booklets were included in the rescoring. Although enough information was captured to calibrate all the fourth-grade items together, this was not the case for the eighth- and eleventh-grade items. There were two discrete groups of items at these two grade levels. Items within the groups were administered to common samples of examinees but there were no common samples for items between these two groups. This prohibited a direct linking via scaling all items simultaneously. As a result, an additional step was required for these two grade levels.

After good starting values were found, as described in Chapter 9, all single-population analyses were done in one run of the NAEP BILOG/PARSCALE computer program without specification of the form of the distribution of proficiencies. All multiple-population analyses, on the other hand, were conducted using fixed normal priors on the proficiency distributions.

Fourth-grade level. For the fourth grade, the generalized partial credit model was used to calibrate, simultaneously, the prompts in the 1984 rescored data and the 1988, 1990, and 1992 data. This required a single run on the NAEP BILOG/PARSCALE computer program specifying four subpopulations, one for each year.

Eighth- and eleventh-grade levels. For these two grade levels, three separate PARSCALE computer runs were conducted: one single-population run on the 1984 original data; and two multiple-population runs, one on each of the two separate groups of prompts discussed above.

15.1.4 Linking the Trend Data and Generating Plausible Values

Once the calibration was completed, as described in section 15.1.3, it was necessary to link the data from the separate scaling runs. Because the scaling was done differently in the fourth grade from the eighth- and eleventh-grades, different linking procedures were required.

Within fourth-grade. No within-grade linking was necessary at this grade level because it was possible to scale all prompts simultaneously across assessment years.

Within eighth- and eleventh-grades. Using a generalization (Muraki & Grima, 1993) of the Stocking-Lord (Stocking & Lord, 1983) procedure, each of the two blocks of prompts at each grade level were linked, separately, to the block of 1984 original prompts. Sampling weights were used in these transformations. The resulting transformation put all items across assessment years on single within-grade scales.

Between-grade linking. At this point in the analyses there were three scales, one at each grade level. The next step was to use the generalized Stocking-Lord procedure to link the within-grade scales so that one common scale could be used for reporting all results. The linking procedure used the eighth-grade scale as a target and separate runs were made to link the grade 4 and the grade 11 scales to that of the eighth grade. Sampling weights were incorporated into this analysis.

The linear transformation constants, *A*, and *B*, produced by the generalized Stocking-Lord procedure are displayed in Table 15-6. Because the age 13/grade 8 sample was the target sample in the transformations, the constants for that sample are one and zero.

Table 15-6
Linear Transformation Constants for the Writing Long-term Trend Between-grade Linking

Age/Grade	A_a	B_a
9/4	1.05	-1.75
13/8	1.00	0.00
17/11	.96	.69

The next step was to generate plausible values to be used in the reporting of all writing trend results, following the procedures described in Chapter 11. Variables based upon the background questions were used for two purposes: as conditioning variables for deriving the plausible values, and to define subgroups for reporting results. Some of these variables are common to all the subject areas; others are specific to the individual areas. Variables used for conditioning and reporting are described in Appendices B and F.

Next, all the data from 1988, 1990, and 1992, but only the rescored data from 1984, were used to set the final metric of the scale. Since the sum of the weights over all individuals in an age/grade is approximately the same (1,000), plausible values for all three grades were combined and the mean and standard deviation were calculated. Note that these statistics were calculated over all five plausible values and all three grade levels. The final step was to transform the scale to the reporting proficiency metric having mean 250 and variance 50.

Overall means and standard deviations in the reporting metric are displayed in Table 15-7.

Table 15-7
Means and Standard Deviations on the Writing Long-term Trend Proficiency Scale

Grade	Assessment Year	All Five Plausible Values	
		Mean	**S.D.**
4	1984	203.8	36.5
	1988	205.7	42.0
	1990	201.7	41.7
	1992	207.1	38.3
8	1984	266.7	29.5
	1988	263.7	32.4
	1990	256.6	37.5
	1992	274.4	36.3
11	1984	289.7	31.8
	1988	291.3	27.9
	1990	287.1	36.5
	1992	287.3	32.0

The variance of proficiency means for each grade was partitioned into the part due to the sampling of students and the part due to the latency of proficiency, θ, as described in Chapter 11. These estimates are given in Table 15-8 (for stability of the estimates, they are based on 100 plausible values). More detailed information is available for gender and race/ethnicity subgroups in Appendix N.

Table 15-8
Estimation Error Variance and Related Coefficients for the Writing Long-term Trend Assessment

Grade	Total Estimation Error Variance	Proportion of Variance Due to...	
		Student Sampling	**Latency of θ**
4	1.40	.43	.57
8	1.81	.61	.39
11	1.11	.45	.55

15.1.5 Other Analyses of Trends in Writing Performance

Between 1988 and 1990 there was a decrease in the eighth-grade mean writing proficiency (263.7 to 256.6). This was followed by an increase (to 274.4) from 1990 to 1992. Although there were somewhat similar trends at the fourth and eleventh grades, they were not as large. A number of analyses were carried out to investigate some of the plausible explanations for these changes in means. The analyses and resulting statistics are reported in Appendix P. It should be noted that the overall means reported in Table 15-7 differ slightly from the means reported in Table 1 of the appendix since the latter are simply the means of the means by item, rather than the overall means that are presented in Table 15-7.

In addition to trends in primary trait scores, trends were measured for the mechanics of writing and for overall writing fluency. Trends in the mechanics of writing at each grade were based on a selected writing prompt given to the grade level in 1984, 1988, 1990, and 1992. The items used for the assessment of the mechanics of writing were "Spaceship" (N0148) for grade 4 and "Recreation Opportunity" (N0003) for grades 8 and 11. All analyses were based on subsamples of approximately 500 responses to each item at each grade and year. Black students were sampled at a higher rate in order to provide sufficient sample size to allow for comparisons in performance between Black and White students. Student weights were adjusted to reflect the oversampling by a poststratification process: For each grade, the students selected for the writing mechanics analysis were categorized by gender and by race/ethnicity (White, Black, Hispanic, other), producing eight cells. The sampling weights of the students within each cell were then multiplied by a poststratification factor computed as a ratio whose denominator was the sum of weights of all students in the cell selected for the mechanics analysis and whose numerator was the sum of the weights of all students in the writing assessment of the specified grade, gender, and race/ethnicity. All papers used in this analysis were scored in 1992; the actual sample sizes are shown in Table 15-9.

Table 15-9
Sample Sizes for Mechanics Scoring

Grade	1984	1988	1990	1992
4	506	484	567	678
8	474	517	601	563
11	522	497	602	566

Two writing items for each grade scored holistically for overall writing fluency. To allow the measurement of trends in overall writing fluency, a sample of responses in the 1984, 1988, and 1990 assessments to the same items were also scored holistically in 1992. Table 15-10 shows the sample sizes for the measurement of trends in the fluency of writing.

Table 15-10
Sample Sizes for Holistic Scoring

Grade	NAEP Item	Sample Size			
		1984	1988	1990	1992
4	Flashlight	2021	611	702	840
	Spaceship	2026	1258	1356	1574
8	Food on Frontier	2235	1341	1502	1308
	Recreation Opportunity	2236	1372	1503	1309
11	Food on Frontier	2372	1212	1399	1436
	Recreation Opportunity	2349	1242	1415	1406

The results of these analyses are reported in *NAEP 1992 Trends in Academic Progress* (Mullis et al., 1994).

15.2 MAIN ASSESSMENT DATA ANALYSIS

The main assessment was administered at grades 4, 8, and 12 using two types of booklets: those containing two prompts, each allotted 25 minutes of writing time, and those containing one prompt, allotting 50 minutes of writing time. The 50-minute prompts were administered only at grades 8 and 12. These double-length writing prompts were included in the 1992 assessment in response to concerns of members of the writing assessment community who argued that the time allotted for responding to a prompt might affect the relationship between performance and writing background (e.g., experience in writing instruction and writing activities). Because of constraints on assessment time, the design could not include booklets containing both 25-minute and 50-minute prompts. The items administered at each grade level are shown in Table 15-11. As described in the writing framework, each prompt used in the main assessment measured one of three purposes of writing: informative, narrative, or persuasive. Table 15-11 also shows the numbers of the blocks containing each writing prompt in each grade.

Table 15-12 provides reliability information for the primary trait scoring of the main assessment writing data. Shown in the table are interrater reliability, percent of exact agreement on scores of the same paper, and percent of agreement within adjacent score categories.

15.2.1 Item Analysis of the Main Assessment Prompts

A standard item analysis for polytomous items, as described in Chapter 9, was conducted on the writing main assessment item data. Item analysis statistics are given in Table 15-13.

15.2.2 Calibration of Main Assessment Prompts and Generation of Plausible Values

Fourth-grade items were calibrated using a straightforward application of the generalized partial credit model. Because of small frequencies, the 5th and 6th scoring categories were combined for three items (W0014, W0015, W0018). Omitted and off-task responses were excluded from the calibration runs.

The calibrations for the eighth- and twelfth-grade data were more complex because of the inclusion of the 50-minute prompts. The fact that 50-minute prompts did not occur in the same booklets as any 25-minute prompts precludes the direct linking of data between these two types of prompts. They could be placed on a common scale only by using a procedure derived by Wang (1993), assuming that length of writing time does not alter the nature of writing proficiency. This procedure will now be described.

Separately for each of the eighth- and twelfth-grade items, the item parameters within each grade were first obtained by calibrating the 25-minute prompts using the generalized partial credit model. Because of small frequencies, three items at the eighth-grade level and two at the twelfth required the collapsing of scoring categories 5 and 6. As in the fourth grade, omitted and off-task responses were treated as missing data.

Table 15-11
Item Number, Purpose of Writing, and Block Number for Each Writing Prompt*

Prompt			Block Number		
Item No.	**Name**	**Type**	**Grade 4**	**Grade 8**	**Grade 12**
25-Minute Blocks					
W0010	School Lunchtime	Informative	3	—	—
W0011	Favorite Story	Informative	5	—	—
W0012	Pet Dinosaur	Narrative	6	—	—
W0013	Magic Balloon	Narrative	8	—	—
W0014	Watch TV	Persuasive	9	—	—
W0015	Space Travelers	Persuasive	11	—	—
W0016	Favorite Object	Informative	4	4	—
W0017	Another Planet	Narrative	7	7	—
W0018	Lengthen School Year	Persuasive	10	10	—
W0020	Performance Review	Informative	—	3	3
W0021	Invention	Informative	—	5	5
W0022	Embarrassing Incident	Narrative	—	6	6
W0023	Grandchildren	Narrative	—	8	8
W0024	Rating Labels	Persuasive	—	9	9
W0025	Drug Search	Persuasive	—	11	11
W0027	Time Capsule	Informative	—	—	4
W0028	Package	Narrative	—	—	7
W0029	Community Service	Persuasive	—	—	10
50-Minute Blocks					
W0019	Dream Car	Narrative	—	13	—
W0026	School Problem	Informative	—	12	12
W0030	History Person	Narrative	—	—	13
W0031	No Pass/No Drive	Persuasive	—	—	14

---

* All prompts have a maximum primary trait score of 6.

Table 15-12
Percentages of Exact Score Agreement and Interrater Reliability for the Primary Trait Scoring of the Writing Main Assessment Items

Items	Grade 4/Age 9				Grade 8/Age 13				Grade 12/Age 17			
	N	Reli-ability	% Agreement	% Adj. Agreement	N	Reli-ability	% Agreement	% Adj. Agreement	N	Reli-ability	% Agreement	% Adj. Agreement
W0010	528	.94	84.9	98.7								
W0011	606	.95	86.5	98.2								
W0012	562	.91	83.6	99.3								
W0013	543	.92	82.3	99.1								
W0014	571	.94	87.2	98.6								
W0015	565	.92	82.7	97.9								
W0016	559	.93	84.8	97.7	621	.93	80.8	99.0				
W0017	595	.90	79.7	98.3	585	.86	73.3	98.8				
W0018	541	.93	82.6	99.5	575	.91	88.9	98.8				
W0019*					617	.91	77.0	97.6				
W0020					625	.94	84.8	98.4	598	.91	83.5	95.8
W0021					595	.87	75.8	98.2	580	.91	83.5	98.3
W0022					600	.86	71.7	98.0	536	.84	69.8	98.0
W0023					604	.86	72.9	98.2	539	.86	73.3	96.1
W0024					597	.94	91.5	99.8	534	.84	83.5	99.1
W0025					596	.92	87.8	99.8	534	.81	82.6	98.5
W0026*					648	.91	79.6	97.5	557	.89	80.6	97.7
W0027									544	.97	92.1	99.8
W0028									540	.81	67.4	97.6
W0029									566	.87	81.3	99.7
W0030*									547	.81	70.6	97.8
W0031*									544	.85	78.5	98.5

---

* 50-minute prompt

Table 15-13

Descriptive Statistics for Writing Prompts, Writing Main Samples

Item	**N**	**Percentages of Examinees in Each Score Category**							**Mean**
		0	**1**	**2**	**3**	**4**	**5**	**6**	
Grade 4									
W0010	1604	7.6	3.1	8.3	42.4	30.8	6.5	1.2	3.36
W0011	1589	8.6	4.0	9.2	45.4	22.0	9.1	1.5	3.31
W0012	1608	4.7	3.8	30.4	37.0	20.9	2.6	0.5	2.89
W0013	1589	5.2	6.4	18.7	40.7	24.9	3.5	0.6	3.02
W0014	1578	14.0	9.1	41.2	28.8	6.1	0.8	0.0	2.40
W0015	1569	10.7	15.0	27.7	31.9	12.8	1.9	0.0	2.54
W0016	1579	10.8	5.9	7.7	43.7	26.4	4.9	0.7	3.21
W0017	1583	6.1	6.6	31.9	35.2	18.1	1.6	0.5	2.76
W0018	1590	10.9	8.4	38.6	33.8	7.9	0.5	0.0	2.48
Grade 8									
W0016	1721	3.4	3.3	2.9	38.9	37.4	12.4	1.7	3.60
W0017	1706	2.4	2.1	14.6	36.2	37.0	7.1	0.6	3.35
W0018	1676	2.0	2.3	22.2	51.5	19.8	2.2	0.0	2.98
W0019	1717	3.3	3.0	27.1	18.6	27.1	14.7	5.9	3.43
W0020	1730	8.6	9.6	3.1	44.6	25.6	8.2	0.2	3.22
W0021	1721	4.7	5.5	17.4	46.5	21.7	4.1	0.2	3.02
W0022	1672	3.2	3.9	28.1	35.1	25.3	4.2	0.1	2.98
W0023	1692	3.0	2.1	31.6	30.7	27.7	3.8	1.0	3.03
W0024	1693	3.3	2.8	29.0	57.5	7.1	0.3	0.0	2.72
W0025	1704	3.7	4.5	33.0	50.8	7.7	0.2	0.0	2.65
W0026	1704	3.3	1.6	4.1	22.6	41.4	23.5	3.4	3.95
Grade 12									
W0020	1539	4.0	5.9	1.4	47.1	24.2	16.6	0.9	3.49
W0021	1542	3.6	6.7	9.2	53.5	21.4	5.2	0.4	3.11
W0022	1551	1.4	2.8	9.7	27.5	42.8	13.6	2.2	3.62
W0023	1540	1.4	2.8	21.8	31.3	29.7	10.9	2.0	3.30
W0024	1541	1.4	2.5	19.6	62.9	12.7	0.8	0.0	2.89
W0025	1424	1.1	2.7	25.7	58.4	11.4	0.7	0.0	2.81
W0026	1487	1.4	0.8	1.6	10.0	40.6	39.6	5.9	4.36
W0027	1537	2.3	4.6	7.5	30.7	35.1	17.4	2.2	3.61
W0028	1570	1.8	3.7	14.4	32.8	37.4	9.1	0.9	3.37
W0029	1520	2.2	3.0	30.4	51.9	11.4	1.1	0.1	2.77
W0030	1540	1.3	2.3	17.4	41.7	23.0	10.6	3.6	3.33
W0031	1514	0.7	2.7	23.4	48.1	22.1	2.7	0.3	3.00

Next, the data and item parameters from the calibration runs were input to the BGROUP program in order to estimate the empirical distribution of the proficiency variable, using 21 conditioning variables for grade 8 and 20 for grade 12. As Wang points out, this approach suffers from lack of model identification. The reason for lack of identification is that for a single item scored on an m-point scale, there are m-1 nondependent pieces of observed data but m parameters to be estimated (one discrimination parameter and m-1 category parameters). This problem of model identification was resolved by employing multiple-group data (i.e., dividing the sample into groups representing subpopulations that are expected to differ in proficiency distributions). By using G groups, assuming invariance of the model over groups, the number of nondependent pieces is increased to $G(m\text{-}1)$, while the number of parameters remains m. Two variables, region and size and type of community, were examined for potential use as grouping variables. After a thorough examination of both variables, region was chosen, resulting in four subpopulations. The 50-minute prompts were next calibrated using the generalized partial credit model, specifying the proficiency distribution estimated in the first step as a fixed (known) distribution, and treating the parameters of the 25-minute prompts as fixed (using the PARSCALE calibration parameter estimates). This was done separately within each of the eighth- and twelfth-grade datasets.

The next step was to use the item parameters and item response data in the BGROUP program to generate five plausible values for each examinee. This was done separately for each grade level and separate runs were made for the 25-minute and the 50-minute prompts at each grade level. The numbers of conditioning variables (Appendix F lists these variables) at the fourth-, eighth-, and twelfth-grades were 105, 207, and 121, respectively.

15.2.3 Linking the Main Assessment Data

After the within-grade scaling described in section 15.2.2 was completed for each of the three grades, the weighted generalized Stocking-Lord procedure was used to place the data for the three grades on the same scale. This was accomplished by linking the fourth-grade scale to that of the eighth grade in one analysis and then linking the twelfth-grade scale to the eighth-grade scale in a second analysis. The procedure is similar to that described for the trend analysis. Using the A and B constants produced by the generalized Stocking-Lord procedure the transformed proficiency variable is $\theta^*_{ia} = A_a\theta_{ia} + B_a$, where θ_{ia} represents a proficiency variable before transformation, i indexes the five plausible values, and a the grade (4, 8, or 12). The constants are displayed in Table 15-14. Those for the eighth-grade scale were 1.0 and 0.0, since the other grades' scales were linked to it.

Table 15-14
Linear Transformation Constants for the Writing Main Between-grade Linking

Grade	A_a	B_a
4	1.12	-1.12
8	1.00	0.00
12	1.07	0.63

In the final step, a linear transformation was made so that the metric for the final scale was set to have a mean of 250 and standard deviation of 50 across the three grades. As in the case of the trend data, this was done by combining all five plausible values for all three grade levels. The resulting mean proficiencies for grades 4, 8, and 12 were 222, 262, and 286, respectively.

The variance of proficiency means for each grade was partitioned into the part due to the sampling of students and the part due to the latency of proficiency, θ, as described in Chapter 11. These estimates are given in Table 15-15 (for stability of the estimates, they are based on 100 plausible values). More detailed information is available for gender and race/ethnicity subgroups in Appendix N.

Table 15-15
Estimation Error Variance and Related Coefficients for the Writing Main Assessment

Grade	**Total Estimation Error Variance**	**Proportion of Variance Due to...**	
		Student Sampling	**Latency of θ**
4	1.01	0.68	0.32
8	0.64	0.65	0.35
12	0.94	0.76	0.24

15.3 WRITING PORTFOLIO DATA ANALYSIS

The writing portfolio study was conducted on a subset of the national writing sample; that is, only age 9/grade 4 and age 13/grade 8 students who were administered booklets 60, 64, 71, 72, and 75 were asked to participate. At age 9/grade 4, the response rate was 91 percent; papers were received from 1,800 of the 1,989 students who were sampled for the school-based writing study. At age 13/grade 8, the response rate was 89 percent; papers were received from 1,895 of the 2,123 sampled students.

The portfolio papers were scored using two sets of criteria. One set consisted of descriptive criteria, the other of evaluative criteria. Both sets were derived from recommendations made by a committee of outside educators with expertise in the field and ETS staff.

Frequency distributions were formed of the results obtained from the descriptive and evaluative scoring. Matches of examinees on the scored portfolio data and the trend writing assessment data were found. Relationships between the descriptive and evaluative scoring and the writing assessment were examined. Results are reported in *Windows into the Classroom: NAEP's 1992 Writing Portfolio Study* (Gentile & Martin-Rehrmann, 1994).

PART III

Statistical Summary of 1992 NAEP Data

Chapter 16

STATISTICAL SUMMARY OF THE 1992 NAEP SAMPLES AND ESTIMATES OF THE PROFICIENCIES OF AMERICAN STUDENTS[1]

Bruce A. Kaplan, Laura S. Jerry, and Phillip T. Leung

Educational Testing Service

The analysis of the 1992 NAEP data has resulted in the production of many thousands of tables presenting estimates of the proficiency of students, and various subgroups of students, in American schools. This chapter gives some selected results from the assessment as well as a statistical summary of the 1992 NAEP samples. The chapter assumes a general familiarity with the structure of NAEP as summarized in the Introduction and overview Chapters 1 and 9.

Three of the many types of NAEP results are presented here:

- results of the instrument development process, including the sizes of the item pools and numbers of booklets;
- results of the sampling process, including the numbers of students in each sample by selected subgroups; and
- results of the parameter estimation process, including estimates of the proficiencies of several populations of students in reading, mathematics, science, and writing.

Interpretive results from the estimates presented here have been reported in the NAEP subject area trend and cross-sectional reports. The 1992 secondary-use data files and user guide (Rogers, Kline, Johnson, Mislevy, & Rust, 1994) are available for those who wish to estimate other parameters of student performance from the NAEP data or to search for possible explanations for the population characteristics that are reported here.

The technical details of the estimation process that underlie these tables are covered in previous parts of this report and not repeated here. A detailed discussion of how to read and use the tables of background and proficiency results is given by Zwick (1987b).

---

[1] Information for various tables in this chapter was provided by Yim Fai Fong, David Freund, Steven Isham, Laura Jerry, Edward Kulick, Michael Narcowich, and Keith Rust.

16.1 MEASUREMENT INSTRUMENTS

For the 1992 assessment, 148 different assessment booklets and questionnaires were printed for age class 9, 157 for age class 13, and 159 for age class 17. These instruments are shown by age level and type in Table 16-1.

The item pool used to develop all main and long-term trend booklets is described in Table 16-2. In general, there are two types of items, cognitive and noncognitive. The cognitive items are developed to measure proficiency in particular subject areas, such as reading and mathematics. Cognitive items may be constructed-response or multiple-choice. The noncognitive items are usually questions about the student's or teacher's backgrounds and perceptions but may also probe other areas, such as school policies or teaching methods. Because many items were used at more than one age class, the total number of items in an item pool is not the sum of the item pools used for the three age classes.

Table 16-3 shows the number of cognitive items in each subject area that were used in the separate samples.

The excluded student, teacher, and school questionnaires contained only noncognitive questions. The number of items in the noncognitive pools is the same as the number of items in the questionnaires. More information about the instruments that were developed is provided in Chapters 2 and 4.

16.2 SAMPLE CHARACTERISTICS

In this section, the characteristics of the final NAEP samples are described. The process by which the sample were selected is discussed in Chapter 3.

In the 1992 main assessment, NAEP contacted 2,250 schools, of which 1,582 contributed data to the assessment. The disposition of these schools is shown in Table 16-4. Some of the schools were unwilling to cooperate; others were believed to be eligible from the sampling frame, but were not. The cooperation rate is calculated as the sum of cooperating schools and the schools that were found to have no eligible students divided by the same sum plus the schools that refused or were from districts that refused to cooperate.

Table 16-4 also shows the number of schools in several categories: region of the country (Northeast, Southeast, Central, West), school governance (public, private, Catholic, Bureau of Indian Affairs, Department of Defense), size and type of community, number of teachers, and number of students.

For the 1992 long-term trend studies, NAEP contacted 1,000 schools, of which 775 contributed data to the various trend assessments. Table 16-5 supplies the same information for the schools assessed for the long-term trend studies that the previous table supplies for the main assessment schools.

The numbers of respondents to the teacher questionnaires are summarized in Table 16-6. The first column in this table includes the number of teachers who responded, by grade and subject area. The second column is the number of students who were not linked to teachers. The third column is the number of students linked to teachers, but not specific classes of these teachers. The last column is the number of students linked to their teachers and their specific classes. (See section 9.2.5 in Chapter 9 for a discussion of teacher/student matching.)

NAEP is administered in units called assessment sessions. If the number of students attending an assessment session is fewer than a predetermined number, the students missing from the session are assigned to a makeup session and then assessed. Table 16-7 shows the number of regular and makeup sessions in 1992 NAEP by age class for the main NAEP and the long-term trend samples. Altogether, 164,414 assessed and excluded students were involved in the 1992 NAEP. The breakdown by age class and by sample is shown in Table 16-8.

Tables 16-9 through 16-21 display the distribution of the students assessed in the main NAEP assessment in several basic categories: gender, racial/ethnic grouping, region of the country, parental education, and size and type of community. There is one table for each main sample taken: reading, mathematics, mathematics estimation, and writing for each of the three age classes, and mathematics with a calculator for age class 9. These tables have four columns:

- eligible by age, which means that the students were in an appropriate age group;
- eligible by grade, which means that the students were in an appropriate grade;
- eligible by age and by grade, which means that the students were of both an appropriate age and appropriate grade; and
- eligible by age or by grade, which is the total number of students for whom data were collected.

Tables 16-22 through 16-27 contain the distribution of students in the same categories by age class for the long-term trend samples. Tables 16-22 to 16-24 contain the distributions for the reading and writing long-term trend samples. Tables 16-25 to 16-27 display the distributions for the mathematics and science long-term trend samples.

Similarly, Tables 16-28 through 16-30 contain the distribution of excluded students by age class for the main samples. Tables 16-31 through 16-33 enumerate the excluded students across the various long-term trend samples.

16.3 POPULATION ESTIMATES

The 1992 NAEP samples were designed for estimating the size and attributes of a number of different populations of students. The estimation procedures use sampling weights, developed by Westat, Inc., that are used in conjunction with the members of the sample (see Chapter 3). In this chapter, all estimates of population parameters use these sampling weights.

Table 16-34 shows the sizes of the various samples and the estimated population sizes by age/grade. The sum of the initial weights for a given sample is an estimate of the number of students who are in the population represented by the sample. In other words, the sum of the initial weights is taken as the estimated population size. In analyses, however, this sum of weights was rescaled to sum to the sample size.

Due to design considerations the main assessment was divided into subsamples, and were administered, and therefore weighted, independently, so that the sum of the initial weights for each subsample estimates the population size. The subsamples were reading, mathematics, mathematics estimation, and writing for all three age classes; and a mathematics calculator bridge for the youngest age class.

Note that the samples for the main assessment and the samples for all three age classes of the reading and writing long-term trend are grade and age samples. The samples for the mathematics and science long-term trend are age-only samples. The sum of the initial weights of the excluded students estimates the number of ineligible students at the respective age/grade levels.

In most cases, the number of students in an age/grade combination is not of interest; a researcher will be interested in estimating the number of students at either a grade or an age level. For the samples that contain both grade- and age-eligible students, an estimate of the total number of students at an age level can be made by summing the initial weights of only the age-eligible students and adding the corresponding sample of age-eligible excluded students' initial weights. An estimate of the total number of students in a grade sample can be made by summing the initial weights of grade-eligible students plus the initial weights of grade-eligible students from the appropriate excluded student sample.

The next group of tables estimates how many students in the main NAEP samples are age-eligible and grade-eligible by age class. Tables 16-35 through 16-47 show how many students at a particular grade level are at, in, or above the modal age for that grade, and how many at a particular age level are at, in, or above the modal grade for that age. Along with the counts from these samples are estimates of the numbers of students in these categories in the population. The standard errors of these estimates and coefficients of variation are also given. (The coefficient of variation of the estimated population size is defined as 100 times its standard error divided by the estimated population size.)

Tables 16-48 through 16-51 contain similar information for the long-term trend booklets, by age level. Where age-only samples are shown, information for all three ages is given in one table, since the partitioning of the sample by modal age groupings provides no added information.

Tables 16-52 to 16-76 show the sizes of the estimated populations of assessable students and the weighted percentages for the NAEP reporting categories of gender, race/ethnicity, region of the country, parents' education level, and size and type of community. The estimated subpopulation percentages for the main NAEP samples are shown in Tables 16-52 through 16-64, separately by age eligibility, grade eligibility, and age/grade eligibility. Tables 16-65 to 16-70 show the same information for the long-term trend samples. In a similar manner, Tables 16-71 to 16-76 show the estimated total population of excluded students and the weighted percentages by demographic subgroups (data about parents' education level is not collected for excluded students and therefore not reported; data about reasons for exclusion are included instead).

Students were assigned proficiency values in a subject area only if they received at least one assessment block in that subject that was scaled. Thus, the sample sizes of students who have proficiency values vary from one subject area to another. Tables 16-77 and 16-78 show the number of students with proficiency values in each subject area by age and grade combinations. Table 16-77 is for the main assessment; Table 16-78 is for the long-term trend assessments.

The rest of the tables in this chapter provide selected proficiency results for students sampled in the 1992 assessment. Tables 16-79 to 16-99 contain percentile information for all NAEP subsamples. These percentile tables are in two forms. For the main assessments, these tables contain population estimates of student proficiencies by grade and by the subpopulations of gender, race/ethnicity, and parents' education level. The information about proficiency includes the mean and standard deviation of each subpopulation as well as the value of the 5th, 10th, 25th, 50th (median), 75th, 90th, and 95th percentiles. Results are shown separately for each subject area. Standard errors of the estimates are included in parentheses. For the long-term trend samples the percentile tables are displayed a little differently, for the total sample and male, female, white, black, and Hispanic subsamples. The reading, mathematics, and science trend samples are age samples; the writing trend sample is a grade-only sample. One column is shown for each trend year. Note that the trend years differ by subject area.

Tables 16-100 through 16-138 contain results for more finely defined subpopulations. For the main assessment three of the major reporting variables—gender, race/ethnicity, and parents' education level—are cross-classified with one another (for example, Table 16-101 cross-classifies gender, race/ethnicity, and parents' education level with the race/ethnicity grouping for fourth-graders in the main reading sample). The data from these and other cross-classifications were used in the creation of the 1992 subject-area reports. Information provided for subpopulations includes the actual sample size (N); the estimated proportion of the population and its standard error (WEIGHTED PCT); the coefficient of variation of the estimated population size (<CV>); the proportion of students in each subpopulation and its standard error; and the average proficiency of the students and its standard error (shown directly below the corresponding proportion). The tables for the long-term trend samples cross the major reporting groups with the various trend years. The statistics reported are the percentages and means with their respective standard errors.

Table 16-1
Measurement Instruments Used in 1992 NAEP

	Age Class		
	9	13	17
Student Assessment Booklets			
Main and Special	62	66	68
Reading	16	20	21
Mathematics	26	26	26
Mathematics Estimation	1	1	1
Mathematics Calculator Bridge	1	0	0
Writing	18	20	21
Long-term Trend	9	9	8
Reading and Writing	6	6	6
Mathematics and Science	3	3	2
Questionnaires	3	4	2
Excluded Student	1	1	1
Teacher	1	2	0
School	1	1	1

Table 16-2
Number of Items Administered, by Age Class

	Age Class			Total Distinct Items
	9	13	17	
Common Background	105	113	147	162
Main	41	41	52	52
Reading and Writing Long-term Trend	64	70	87	99
Mathematics and Science Long-term Trend	20	21	28	33
Reading Background	59	71	107	110
Main	19	29	29	29
Long-term Trend	40	42	78	81
Reading Cognitive	190	242	240	488
Main	85	134	144	295
Long-term Trend	105	108	96	193
Mathematics Background	26	57	78	101
Main	23	28	39	52
Long-term Trend	3	29	39	49
Mathematics Cognitive	261	333	338	637
Main	193	235	244	454
Long-term Trend	68	98	94	184
Science Long-term Trend				
Background	16	29	29	45
Cognitive	63	83	82	180
Writing Background	74	96	96	109
Main	21	31	31	32
Long-term Trend	53	65	65	77
Writing Cognitive	15	17	18	34
Main	9	11	12	22
Long-term Trend	6	6	6	12
Excluded Student Questionnaire	67	67	67	67
Teacher Questionnaire	164	171	0	244
School Questionnaire	91	90	88	141
Total Items	1131	1369	1290	2318

Table 16-3
Number of Cognitive Items Administered by Sample and Age Class

	Age Class			
	9	13	17	Total
Main and Special				
Reading	85	134	144	295
Mathematics	153	183	180	367
Mathematics Estimation	54	73	86	110
Mathematics Calculator Bridge	37	0	0	37
Writing	9	11	12	22
Long-term Trend				
Reading	105	108	96	193
Mathematics	68	98	94	184
Writing	6	6	6	12
Science	63	83	82	180
Total Cognitive Items*	529	675	678	1339

---

* Because many items were used at more than one age class and/or for more than one sample, the total number of cognitive items is not equal to the total number of distinct items used for the three age classess and across the samples. Item counts in this table are for all items presented, but not necessarily used in scaling. Therefore, the number of items shown here may not agree with item counts reported in other chapters.

Table 16-4
Characteristics* of Schools in Main Samples

	Age/Grade			
	9/4	13/8	17/12	Total
Total original sample	669	824	757	2250
Cooperating	523	583	464	1570
NELS	0	5	71	76
Out-of-range or closed	50	84	81	215
No eligibles enrolled	12	44	29	85
District refused	33	25	34	92
School refused	51	83	78	212
Cooperation rate	86%	85%	81%	84%
Cooperating replacements for refusals	4	4	4	12
Totals				
Cooperating schools	527	587	468	1582
Completed questionnaires	517	553	439	1509
Region				
Northeast	110	136	90	336
Southeast	121	127	128	376
Central	160	181	116	457
West	136	143	134	413
School type				
Public	319	303	321	943
Private	74	109	94	277
Catholic	133	175	53	361
Bureau of Indian Affairs	0	0	0	0
Department of Defense	1	0	0	1
Size and type of community				
Rural	58	59	67	184
Disadvantaged urban	54	57	60	171
Advantaged urban	75	75	51	201
Big city	50	74	47	171
Fringe	63	93	52	208
Medium city	90	92	51	233
Small place	137	137	140	414

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-4 (continued)
Characteristics of Schools in Main Samples

	Age/Grade			
	9/4	13/8	17/12	Total
Number of Teachers				
Unclassified	0	0	0	0
1 - 4	13	18	1	32
5 - 9	80	86	32	198
10 - 19	177	164	67	408
20 - 49	234	213	150	597
50 - 74	10	69	89	168
75 - 99	3	20	60	83
100 +	0	8	56	64
Missing	10	9	13	32
Number of Students				
Unclassified	0	0	0	0
1 - 99	31	29	21	81
100 - 299	184	212	96	492
300 - 499	145	115	71	331
500 - 749	111	95	48	254
750 - 999	32	67	53	152
1000 - 1499	12	45	77	134
1500 +	2	15	89	106
Missing	10	9	13	32

Table 16-5
Characteristics* of Schools in Long-term Trend Samples

	Age Class			
	9	13	17	Total
Total original sample	363	327	310	1000
Cooperating	303	248	213	764
NELS	0	2	40	42
Out-of-range or closed	6	4	3	13
No eligibles enrolled	9	24	4	37
District refused	31	35	25	91
School refused	14	14	25	53
Cooperation rate	87	85	81	85
Cooperating replacements for refusals	4	3	4	11
Totals				
Cooperating schools	307	251	217	775
Completed questionnaires	305	249	212	766
Region				
Northeast	61	53	46	160
Southeast	68	56	53	177
Central	97	79	63	239
West	81	63	55	199
School type				
Public	251	196	196	643
Private	20	22	9	51
Catholic	36	33	11	80
Bureau of Indian Affairs	0	0	1	1
Department of Defense	0	0	0	0
Size and type of community				
Rural	38	32	30	100
Disadvantaged urban	28	26	29	83
Advantaged urban	29	29	21	79
Big city	30	33	10	73
Fringe	46	36	42	124
Medium city	47	40	21	108
Small place	89	55	64	208

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-5 (continued)
Characteristics of Schools in Long-term Trend Samples

	Age Class			
	9	13	17	Total
Number of Teachers				
Unclassified	0	0	0	0
1 - 4	7	3	1	11
5 - 9	24	21	4	49
10 - 19	90	43	21	154
20 - 49	174	123	67	364
50 - 74	9	48	41	98
75 - 99	2	8	39	49
100 +	0	5	41	46
Missing	1	0	3	4
Number of Students				
Unclassified	0	0	0	0
1 - 99	10	6	4	20
100 - 299	75	56	25	156
300 - 499	103	54	21	178
500 - 749	72	63	34	169
750 - 999	39	39	24	102
1000 - 1499	7	29	42	78
1500 +	0	4	64	68
Missing	1	0	3	4

Table 16-6
Numbers of Responses to Teacher Questionnaires and Students Matched with Teacher Data

	Number of Teachers Responding	Number of Students with No Match	Number of Students with Partial Match	Number of Students with Complete Match
Grade 4 Reading	894	576	1176	4562
Grade 4 Mathematics	968	1257	1220	6366
Grade 8 Mathematics	835	809	749	6105
Grade 8 Writing	958	1045	1262	8805

Table 16-7
Numbers of Assessment Sessions by Sample, Type of Session, and Age Class

	Age Class			
	9	13	17	Total
Main Sample				
Regular sessions	1422	1339	1288	4049
Makeup sessions	1	11	72	84
Reading and Writing Long-term Trend Sample				
Regular sessions	310	232	264	806
Makeup sessions	1	0	6	7
Mathematics and Science Long-term Trend Sample				
Regular sessions	471	346	279	1096
Makeup sessions	1	5	10	16
Total				
Regular sessions	2203	1917	1831	5951
Makeup sessions	3	16	88	107

Table 16-8
Number of Students Assessed and Excluded by Sample and Age Class

	Age Class 9	Age Class 13	Age Class 17	Total
Assessed				
Main	31672	42591	42556	116819
Reading and Writing Long-term Trend	7062	5514	5569	18145
Mathematics and Science Long-term Trend	7335	5909	4359	17603
Excluded				
Main	2713	3325	2835	8873
Long-term Trend	1303	867	804	2974
Total	50085	58206	56123	164414

Table 16-9
Number of Students in Reading Main Sample by Type of Eligibility and Subgroup Classification*
Age 9/grade 4

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Total	5966	6314	3864	8416
Sex				
Male	2950	3171	1783	4338
Female	3016	3143	2081	4078
Race/Ethnicity				
White	3724	3917	2449	5192
Black	863	1013	558	1318
Hispanic	1051	1044	641	1454
Other	328	340	216	452
Region				
Northeast	1147	1228	845	1530
Southeast	1563	1649	919	2293
Central	1539	1599	978	2160
West	1717	1838	1122	2433
Parents' Education				
Less than high school	236	291	139	388
High school	742	769	460	1051
Greater than high school	432	529	322	639
Graduated college	2276	2481	1574	3183
Unknown	2259	2228	1359	3128
Size and Type of Community				
Rural	571	608	358	821
Disadvantaged urban	673	742	429	986
Advantaged urban	684	693	476	901
Big city	681	713	473	921
Fringe	937	1004	668	1273
Medium city	879	951	540	1290
Small places	1541	1603	920	2224

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-10
Number of Students in Mathematics Main Sample by Type of Eligibility and Subgroup Classification*
Age 9/grade 4

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Total	6597	7176	4359	9414
Sex				
Male	3204	3553	1972	4785
Female	3393	3623	2387	4629
Race/Ethnicity				
White	4031	4381	2699	5713
Black	1050	1192	690	1552
Hispanic	1141	1182	703	1620
Other	375	421	267	529
Region				
Northeast	1329	1396	1009	1716
Southeast	1822	2038	1149	2711
Central	1673	1819	1060	2432
West	1773	1923	1141	2555
Parents' Education				
Less than high school	247	298	138	407
High school	762	852	500	1114
Greater than high school	408	522	292	638
Graduated college	2549	2880	1810	3619
Unknown	2622	2616	1617	3621
Size and Type of Community				
Rural	620	706	394	932
Disadvantaged urban	751	825	462	1114
Advantaged urban	851	891	607	1135
Big city	733	765	540	958
Fringe	928	1021	668	1281
Medium city	1098	1173	650	1621
Small places	1616	1795	1038	2373

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-11
Number of Students in Mathematics Estimation Sample by Type of Eligibility and Subgroup Classification*
Age 9/grade 4

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Total	1418	1562	926	2054
Sex				
Male	682	764	421	1025
Female	736	798	505	1029
Race/Ethnicity				
White	902	991	595	1298
Black	226	260	158	328
Hispanic	233	241	131	343
Other	57	70	42	85
Region				
Northeast	277	294	207	364
Southeast	410	475	249	636
Central	392	429	242	579
West	339	364	228	475
Parents' Education				
Less than high school	54	66	31	89
High school	181	206	117	270
Greater than high school	89	127	73	143
Graduated college	558	625	384	799
Unknown	534	533	319	748
Size and Type of Community				
Rural	102	122	57	167
Disadvantaged urban	161	182	97	246
Advantaged urban	191	184	141	234
Big city	94	113	71	136
Fringe	215	242	151	306
Medium city	257	291	166	382
Small places	398	428	243	583

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-12
Number of Students in Mathematics Calculator Sample by Type of Eligibility and Subgroup Classification*
Age 9/grade 4

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Total	1594	1667	1025	2236
Sex				
Male	786	828	470	1144
Female	808	839	555	1092
Race/Ethnicity				
White	1005	1049	649	1405
Black	252	293	178	367
Hispanic	254	250	149	355
Other	83	75	49	109
Region				
Northeast	321	335	238	418
Southeast	435	471	266	640
Central	381	400	238	543
West	457	461	283	635
Parents' Education				
Less than high school	74	91	49	116
High school	199	214	126	287
Greater than high school	93	107	61	139
Graduated college	618	655	421	852
Unknown	610	598	368	840
Size and Type of Community				
Rural	140	154	73	221
Disadvantaged urban	181	182	106	257
Advantaged urban	244	258	181	321
Big city	152	144	106	190
Fringe	186	189	128	247
Medium city	263	288	163	388
Small places	428	452	268	612

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-13
Number of Students in Writing Main Sample by Type of Eligibility and Subgroup Classification*
Age 9/grade 4

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Total	6752	7166	4366	9552
Sex				
Male	3263	3483	1930	4816
Female	3489	3683	2436	4736
Race/Ethnicity				
White	4225	4420	2735	5910
Black	1067	1205	714	1558
Hispanic	1059	1119	644	1534
Other	401	422	273	550
Region				
Northeast	1381	1473	1023	1831
Southeast	1821	1944	1112	2653
Central	1706	1741	1035	2412
West	1844	2008	1196	2656
Parents' Education				
Less than high school	264	305	149	420
High school	881	958	581	1258
Greater than high school	484	619	367	736
Graduated college	2618	2856	1804	3670
Unknown	2485	2407	1452	3440
Size and Type of Community				
Rural	580	602	328	854
Disadvantaged urban	772	853	484	1141
Advantaged urban	860	873	626	1107
Big city	656	698	469	885
Fringe	933	1015	657	1291
Medium city	1090	1213	645	1658
Small places	1861	1912	1157	2616

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-14
Number of Students in Reading Main Sample by Type of Eligibility and Subgroup Classification*
Age 13/grade 8

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Total	10579	11184	6821	14942
Sex				
Male	5171	5561	3078	7654
Female	5408	5623	3743	7288
Race/Ethnicity				
White	6609	7069	4424	9254
Black	1788	1867	1071	2584
Hispanic	1602	1655	940	2317
Other	580	593	386	787
Region				
Northeast	2322	2370	1638	3054
Southeast	2655	2904	1644	3915
Central	2753	2906	1734	3925
West	2849	3004	1805	4048
Parents' Education				
Less than high school	757	871	405	1223
High school	2340	2513	1425	3428
Greater than high school	1910	2146	1352	2704
Graduated college	4530	4686	3127	6089
Unknown	1026	940	500	1466
Size and Type of Community				
Rural	696	734	415	1015
Disadvantaged urban	1387	1486	866	2007
Advantaged urban	1080	1182	841	1421
Big city	1401	1466	903	1964
Fringe	1559	1620	1132	2047
Medium city	1744	1767	929	2582
Small places	2712	2929	1735	3906

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-15
Number of Students in Mathematics Main Sample by Type of Eligibility and Subgroup Classification*
Age 13/grade 8

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Total	7290	7663	4662	10291
Sex				
Male	3579	3820	2100	5299
Female	3711	3843	2562	4992
Race/Ethnicity				
White	4672	4903	3100	6475
Black	1156	1195	672	1679
Hispanic	1077	1120	621	1576
Other	385	445	269	561
Region				
Northeast	1526	1573	1074	2025
Southeast	1814	1908	1110	2612
Central	1890	1921	1163	2648
West	2060	2261	1315	3006
Parents' Education				
Less than high school	508	563	271	800
High school	1518	1679	924	2273
Greater than high school	1329	1382	903	1808
Graduated college	3133	3290	2160	4263
Unknown	772	726	394	1104
Size and Type of Community				
Rural	531	581	319	793
Disadvantaged urban	803	841	479	1165
Advantaged urban	821	874	626	1069
Big city	946	986	597	1335
Fringe	1112	1177	798	1491
Medium city	1161	1197	673	1685
Small places	1916	2007	1170	2753

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-16
Number of Students in Mathematics Estimation Sample by Type of Eligibility and Subgroup Classification*
Age 13/grade 8

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Total	1664	1769	1017	2416
Sex				
Male	857	933	485	1305
Female	807	836	532	1111
Race/Ethnicity				
White	1013	1091	643	1461
Black	270	311	168	413
Hispanic	299	291	155	435
Other	82	76	51	107
Region				
Northeast	348	378	229	497
Southeast	401	461	246	616
Central	475	499	286	688
West	440	431	256	615
Parents' Education				
Less than high school	121	150	61	210
High school	378	417	218	577
Greater than high school	271	313	186	398
Graduated college	694	726	455	965
Unknown	197	163	97	263
Size and Type of Community				
Rural	107	125	72	160
Disadvantaged urban	241	245	140	346
Advantaged urban	249	284	195	338
Big city	140	160	89	211
Fringe	238	239	161	316
Medium city	345	324	165	504
Small places	344	392	195	541

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-17
Number of Students in Writing Main Sample by Type of Eligibility and Subgroup Classification*
Age 13/grade 8

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Total	10633	11112	6803	14942
Sex				
Male	5288	5583	3129	7742
Female	5345	5529	3674	7200
Race/Ethnicity				
White	6743	7083	4435	9391
Black	1688	1785	1026	2447
Hispanic	1607	1621	926	2302
Other	595	623	416	802
Region				
Northeast	2356	2416	1676	3096
Southeast	2535	2755	1536	3754
Central	2839	2917	1752	4004
West	2903	3024	1839	4088
Parents' Education				
Less than high school	728	813	386	1155
High school	2347	2534	1413	3468
Greater than high school	1939	2133	1375	2697
Graduated college	4561	4670	3096	6135
Unknown	1028	936	520	1444
Size and Type of Community				
Rural	801	828	466	1163
Disadvantaged urban	1275	1350	757	1868
Advantaged urban	1087	1196	816	1467
Big city	1389	1437	895	1931
Fringe	1596	1660	1175	2081
Medium city	1769	1793	992	2570
Small places	2716	2848	1702	3862

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-18
Number of Students in Reading Main Sample by Type of Eligibility and Subgroup Classification*
Age 17/grade 12

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Total	11681	11355	7722	15314
Sex				
Male	5687	5432	3426	7693
Female	5994	5923	4296	7621
Race/Ethnicity				
White	7995	7759	5489	10265
Black	1953	1841	1161	2633
Hispanic	1136	1151	686	1601
Other	597	604	386	815
Region				
Northeast	2634	2646	1941	3339
Southeast	3342	3121	2114	4349
Central	2743	2686	1806	3623
West	2962	2902	1861	4003
Parents' Education				
Less than high school	877	898	457	1318
High school	2555	2339	1516	3378
Greater than high school	3010	2951	2124	3837
Graduated college	4920	4849	3478	6291
Unknown	291	286	130	447
Size and Type of Community				
Rural	967	913	582	1298
Disadvantaged urban	1678	1524	950	2252
Advantaged urban	1491	1505	1148	1848
Big city	1293	1326	919	1700
Fringe	1575	1563	1152	1986
Medium city	1163	1111	728	1546
Small places	3514	3413	2243	4684

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-19
Number of Students in Mathematics Main Sample by Type of Eligibility and Subgroup Classification*
Age 17/grade 12

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Total	7328	6973	4802	9499
Sex				
Male	3509	3290	2086	4713
Female	3819	3683	2716	4786
Race/Ethnicity				
White	4985	4738	3418	6305
Black	1257	1122	744	1635
Hispanic	711	734	402	1043
Other	375	379	238	516
Region				
Northeast	1652	1641	1211	2082
Southeast	2150	1934	1339	2745
Central	1646	1569	1055	2160
West	1880	1829	1197	2512
Parents' Education				
Less than high school	523	495	268	750
High school	1562	1429	901	2090
Greater than high school	1889	1793	1310	2372
Graduated college	3125	3048	2207	3966
Unknown	203	181	98	286
Size and Type of Community				
Rural	793	738	469	1062
Disadvantaged urban	1045	906	611	1340
Advantaged urban	916	906	691	1131
Big city	892	886	626	1152
Fringe	881	884	648	1117
Medium city	766	714	454	1026
Small places	2035	1939	1303	2671

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-20
Number of Students in Mathematics Estimation Sample by Type of Eligibility and Subgroup Classification*
Age 17/grade 12

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Total	1575	1526	1027	2074
Sex				
Male	802	799	488	1113
Female	773	727	539	961
Race/Ethnicity				
White	1054	1021	732	1343
Black	277	260	154	383
Hispanic	156	143	79	220
Other	88	102	62	128
Region				
Northeast	324	348	236	436
Southeast	514	464	314	664
Central	344	358	248	454
West	393	356	229	520
Parents' Education				
Less than high school	119	105	56	168
High school	360	328	206	482
Greater than high school	396	375	262	509
Graduated college	653	665	480	838
Unknown	40	45	18	67
Size and Type of Community				
Rural	72	73	49	96
Disadvantaged urban	258	236	135	359
Advantaged urban	208	206	164	250
Big city	187	181	119	249
Fringe	187	208	149	246
Medium city	149	149	96	202
Small places	514	473	315	672

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-21
Number of Students in Writing Main Sample by Type of Eligibility and Subgroup Classification*
Ag 17/grade 12

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Total	12064	11532	7927	15669
Sex				
Male	5842	5515	3481	7876
Female	6222	6017	4446	7793
Race/Ethnicity				
White	8229	7821	5619	10431
Black	2095	1949	1228	2816
Hispanic	1156	1159	697	1618
Other	584	603	383	804
Region				
Northeast	2692	2670	1917	3445
Southeast	3553	3207	2241	4519
Central	2812	2730	1874	3668
West	3007	2925	1895	4037
Parents' Education				
Less than high school	904	875	486	1293
High school	2595	2420	1579	3436
Greater than high school	3093	2993	2130	3956
Graduated college	5131	4952	3584	6499
Unknown	310	265	132	443
Size and Type of Community				
Rural	1020	950	623	1347
Disadvantaged urban	1758	1582	1020	2320
Advantaged urban	1606	1572	1218	1960
Big city	1364	1398	946	1816
Fringe	1624	1580	1202	2002
Medium city	1222	1132	728	1626
Small places	3470	3318	2190	4598

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-22
Number of Students in Reading and Writing Long-term Trend Sample
by Type of Eligibility and Subgroup Classification*
Age 9/grade 4

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Total	4944	5176	3058	7062
Sex				
Male	2491	2578	1399	3670
Female	2453	2598	1659	3392
Race/Ethnicity				
White	3277	3445	2043	4679
Black	679	764	419	1024
Hispanic	710	687	407	990
Other	278	280	189	369
Region				
Northeast	973	1058	713	1318
Southeast	1068	1160	606	1622
Central	1507	1538	910	2135
West	1396	1420	829	1987
Parents' Education				
Less than high school	238	245	131	352
High school	758	845	447	1156
Greater than high school	241	278	147	372
Graduated college	1962	2147	1331	2778
Unknown	1708	1645	994	2359
Size and Type of Community				
Rural	537	583	292	828
Disadvantaged urban	512	518	300	730
Advantaged urban	464	479	305	638
Big city	489	549	363	675
Fringe	783	760	486	1057
Medium city	836	853	499	1190
Small places	1323	1434	813	1944

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-23
Number of Students in Reading and Writing Long-term Trend Sample by Type of Eligibility and Subgroup Classification*
Age 13/grade 8

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Total	3965	4103	2554	5514
Sex				
Male	1904	1984	1108	2780
Female	2061	2119	1446	2734
Race/Ethnicity				
White	2774	2922	1818	3878
Black	568	576	350	794
Hispanic	411	396	239	568
Other	212	209	147	274
Region				
Northeast	827	845	608	1064
Southeast	905	949	523	1331
Central	1110	1127	704	1533
West	1123	1182	719	1586
Parents' Education				
Less than high school	232	274	121	385
High school	1106	1166	679	1593
Greater than high school	467	505	345	627
Graduated college	1795	1829	1226	2398
Unknown	352	321	179	494
Size and Type of Community				
Rural	370	417	201	586
Disadvantaged urban	403	358	226	535
Advantaged urban	498	521	403	616
Big city	503	502	342	663
Fringe	643	632	418	857
Medium city	713	751	422	1042
Small places	835	922	542	1215

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-24
Number of Students in Reading and Writing Long-term Trend Sample by Type of Eligibility and Subgroup Classification*
Age 17/grade 11

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Total	4447	4339	3217	5569
Sex				
Male	2282	2214	1561	2935
Female	2165	2125	1656	2634
Race/Ethnicity				
White	3375	3304	2577	4102
Black	551	511	304	758
Hispanic	311	302	195	418
Other	210	222	141	291
Region				
Northeast	981	969	725	1225
Southeast	1077	1006	679	1404
Central	1327	1348	1063	1612
West	1062	1016	750	1328
Parents' Education				
Less than high school	295	289	175	409
High school	1258	1160	844	1574
Greater than high school	883	890	684	1089
Graduated college	1883	1892	1451	2324
Unknown	119	102	58	163
Size and Type of Community				
Rural	428	416	308	536
Disadvantaged urban	485	458	279	664
Advantaged urban	461	441	337	565
Big city	197	202	150	249
Fringe	980	973	730	1223
Medium city	446	432	328	550
Small places	1450	1417	1085	1782

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-25
Number of Students in Mathematics and Science Long-term Trend Sample by Type of Eligibility and Subgroup Classification*
Age 9

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Total	7335	4569	4569	7335
Sex				
Male	3581	2081	2081	3581
Female	3754	2488	2488	3754
Race/Ethnicity				
White	4829	3124	3124	4829
Black	966	587	587	966
Hispanic	1221	650	650	1221
Other	319	208	208	319
Region				
Northeast	1402	1015	1015	1402
Southeast	1700	966	966	1700
Central	2301	1386	1386	2301
West	1932	1202	1202	1932
Parents' Education				
Less than high school	321	180	180	321
High school	1015	586	586	1015
Greater than high school	568	389	389	568
Graduated college	2978	1955	1955	2978
Unknown	2423	1443	1443	2423
Size and Type of Community				
Rural	682	367	367	682
Disadvantaged urban	652	375	375	652
Advantaged urban	757	536	536	757
Big city	771	567	567	771
Fringe	1237	796	796	1237
Medium city	1130	663	663	1130
Small places	2106	1265	1265	2106

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-26
Number of Students in Mathematics and Science Long-term Trend Sample by Type of Eligibility and Subgroup Classification*
Age 13

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Total	5909	3813	3813	5909
Sex				
Male	2926	1707	1707	2926
Female	2983	2106	2106	2983
Race/Ethnicity				
White	4149	2740	2740	4149
Black	810	491	491	810
Hispanic	645	360	360	645
Other	305	222	222	305
Region				
Northeast	1271	915	915	1271
Southeast	1404	879	879	1404
Central	1710	1063	1063	1710
West	1524	956	956	1524
Parents' Education				
Less than high school	350	178	178	350
High school	1359	821	821	1359
Greater than high school	1083	776	776	1083
Graduated college	2613	1780	1780	2613
Unknown	489	250	250	489
Size and Type of Community				
Rural	562	303	303	562
Disadvantaged urban	595	328	328	595
Advantaged urban	669	517	517	669
Big city	847	565	565	847
Fringe	908	637	637	908
Medium city	1025	627	627	1025
Small places	1303	836	836	1303

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-27
Number of Students in Mathematics and Science Long-term Trend Sample by Type of Eligibility and Subgroup Classification*
Age 17

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Total	4359	3154	3154	4359
Sex				
Male	2168	1482	1482	2168
Female	2191	1672	1672	2191
Race/Ethnicity				
White	3295	2508	2508	3295
Black	498	291	291	498
Hispanic	366	230	230	366
Other	200	125	125	200
Region				
Northeast	982	722	722	982
Southeast	1075	689	689	1075
Central	1260	994	994	1260
West	1042	749	749	1042
Parents' Education				
Less than high school	302	171	171	302
High school	941	637	637	941
Greater than high school	1107	829	829	1107
Graduated college	1900	1460	1460	1900
Unknown	99	50	50	99
Size and Type of Community				
Rural	601	446	446	601
Disadvantaged urban	490	292	292	490
Advantaged urban	474	347	347	474
Big city	193	142	142	193
Fringe	896	673	673	896
Medium city	406	301	301	406
Small places	1299	953	953	1299

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-28
Number of Excluded Students in Main Sample by Type of Eligibility and Subgroup Classification*
Age 9/grade 4

	Eligible by			
	Age	Grade	Age & Grade	Total
Total	1682	1773	742	2713
Sex				
Male	1002	1079	407	1674
Female	680	694	335	1039
Race/Ethnicity				
White	550	569	185	934
Black	247	237	84	400
Hispanic	778	853	417	1214
Other	107	114	56	165
Region				
Northeast	276	300	134	442
Southeast	388	420	115	693
Central	243	220	75	388
West	775	833	418	1190
Size and Type of Community				
Rural	105	131	46	190
Disadvantaged urban	447	423	228	642
Advantaged urban	71	71	33	109
Big city	161	187	71	277
Fringe	297	302	138	461
Medium city	325	382	127	580
Small places	276	277	99	454
Reason for Exclusion				
Disability	816	899	292	1423
Limited English Proficiency	696	743	406	1033
Both disability and LEP	81	82	17	146
Nonreader	41	25	16	50
Other	48	24	11	61

Table 16-29
Number of Excluded Students in Main Sample by Type of Eligibility and Subgroup Classification*
Age 13/grade 8

	Eligible by			
	Age	Grade	Age & Grade	Total
Total	1996	2030	701	3325
Sex				
Male	1250	1285	406	2129
Female	746	745	295	1196
Race/Ethnicity				
White	749	820	241	1328
Black	428	328	105	651
Hispanic	690	737	288	1139
Other	129	145	67	207
Region				
Northeast	406	368	146	628
Southeast	355	372	80	647
Central	409	433	124	718
West	826	857	351	1332
Size and Type of Community				
Rural	74	72	12	134
Disadvantaged urban	488	476	200	764
Advantaged urban	111	107	52	166
Big city	209	184	55	338
Fringe	426	419	191	654
Medium city	363	361	106	618
Small places	325	411	85	651
Reason for Exclusion				
Disability	1221	1215	327	2109
Limited English Proficiency	655	709	330	1034
Both disability and LEP	82	70	29	123
Nonreader	16	13	6	23
Other	19	20	8	31

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-30
Number of Excluded Students in Main Sample by Type of Eligibility and Subgroup Classification*
Age 17/grade 12

	Eligible by			
	Age	Grade	Age & Grade	Total
Total	1804	1422	391	2835
Sex				
Male	1115	868	227	1756
Female	689	554	164	1079
Race/Ethnicity				
White	730	654	190	1194
Black	398	324	80	642
Hispanic	476	277	82	671
Other	200	167	39	328
Region				
Northeast	402	319	84	637
Southeast	401	321	68	654
Central	342	284	90	536
West	659	498	149	1008
Size and Type of Community				
Rural	139	136	36	239
Disadvantaged urban	479	337	93	723
Advantaged urban	109	96	40	165
Big city	151	95	23	223
Fringe	208	172	41	339
Medium city	289	215	68	436
Small places	429	371	90	710
Reason for Exclusion				
Disability	1158	995	251	1902
Limited English Proficiency	550	356	114	792
Both disability and LEP	62	48	11	99
Nonreader	10	7	4	13
Other	22	16	11	27

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-31
Number of Excluded Students in Long-term Trend Samples by Type of Eligibility and Subgroup Classification*
Age 9/grade 4

	Eligible by			
	Age	Grade	Age & Grade	Total
Total	975	729	401	1303
Sex				
Male	612	443	240	815
Female	363	286	161	488
Race/Ethnicity				
White	398	264	122	540
Black	154	117	58	213
Hispanic	272	222	127	367
Other	151	126	94	183
Region				
Northeast	242	172	111	303
Southeast	223	148	49	322
Central	126	108	51	183
West	384	301	190	495
Size and Type of Community				
Rural	82	55	20	117
Disadvantaged urban	208	165	94	279
Advantaged urban	89	82	56	115
Big city	88	58	36	110
Fringe	167	140	92	215
Medium city	167	121	60	228
Small places	174	108	43	239
Reason for Exclusion				
Disability	538	400	179	759
Limited English Proficiency	302	239	166	375
Both disability and LEP	56	41	27	70
Nonreader	49	32	20	61
Other	30	17	9	38

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-32
Number of Excluded Students in Long-term Trend Samples by Type of Eligibility and Subgroup Classification*
Age 13/grade 8

	Eligible by			
	Age	Grade	Age & Grade	Total
Total	663	419	215	867
Sex				
Male	415	267	133	549
Female	248	152	82	318
Race/Ethnicity				
White	256	164	79	341
Black	121	60	26	155
Hispanic	181	137	77	241
Other	105	58	33	130
Region				
Northeast	148	88	55	181
Southeast	172	90	34	228
Central	123	79	40	162
West	220	162	86	296
Size and Type of Community				
Rural	45	28	9	64
Disadvantaged urban	140	87	43	184
Advantaged urban	67	33	23	77
Big city	77	36	21	92
Fringe	122	85	47	160
Medium city	124	90	45	169
Small places	88	60	27	121
Reason for Exclusion				
Disability	369	236	107	498
Limited English Proficiency	205	133	86	252
Both disability and LEP	22	16	4	34
Nonreader	24	12	5	31
Other	43	21	13	51

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-33
Number of Excluded Students in Long-term Trend Samples by Type of Eligibility and Subgroup Classification*
Age 17/grade 11

	Eligible by			
	Age	Grade	Age & Grade	Total
Total	506	336	163	804
Sex				
Male	338	206	100	513
Female	168	130	63	291
Race/Ethnicity				
White	252	177	91	377
Black	90	49	26	141
Hispanic	88	48	24	160
Other	76	62	22	126
Region				
Northeast	107	87	45	188
Southeast	136	59	23	196
Central	111	85	44	171
West	152	105	51	249
Size and Type of Community				
Rural	54	36	16	78
Disadvantaged urban	118	56	28	202
Advantaged urban	57	53	35	88
Big city	17	9	3	28
Fringe	96	79	38	144
Medium city	35	35	9	83
Small places	129	68	34	181
Reason for Exclusion				
Disability	354	235	115	544
Limited English Proficiency	104	68	30	189
Both disability and LEP	20	15	6	37
Nonreader	14	13	8	19
Other	14	4	4	14

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-34
Number of Students by Sample and Age Class

	— Age Class 9 —		— Age Class 13 —		— Age Class 17 —	
	Total	Sum of Weights	Total	Sum of Weights	Total	Sum of Weights
Main Assessment						
Reading	8416	4912046	14942	4652309	15314	3605818
Mathematics	9414	4922681	10291	4649680	9499	3596526
Mathematic Estimation	2054	4887664	2416	4687928	2074	3610124
Mathematics Calculator Bridge	2236	4931884	0	0	0	0
Writing	9552	4909219	14942	4663137	15669	3615273
Long-term Trend Assessment						
Reading and Writing	7062	4999070	5514	4712846	5569	4064126
Mathematics and Science	7335	3362703	5909	3181507	4359	2936756
Excluded Students						
Main	2713	383975	3325	357725	2835	216504
Long-term Trend	1303	396347	867	333967	804	292533

Table 16-35
Number of Students Assessed in Reading Main Sample
Age 9/grade 4

	Grade			
	< 4	= 4	> 4	Total
AGE < 9				
Unweighted N	0	33	0	33
Estimated population size	0	20670	0	20670
Standard error	0	3341	0	3341
Coefficient of variation*	0.00	16.16	0.00	16.16
AGE = 9				
Unweighted N	2091	3864	11	5966
Estimated population size	1441007	2028390	6738	3476135
Standard error	14470	31700	2510	35106
Coefficient of variation*	1.00	1.56	37.25	1.01
AGE > 9				
Unweighted N	0	2417	0	2417
Estimated population size	0	1415241	0	1415241
Standard error	0	34047	0	34047
Coefficient of variation*	0.00	2.41	0.00	2.41
AGE TOTAL				
Unweighted N	2091	6314	11	8416
Estimated population size	1441007	3464301	6738	4912046
Standard error	14470	16618	2510	28689
Coefficient of variation*	1.00	0.48	37.25	0.58

---

* The coefficient of variation of the estimated population is defined as 100 times its standard error divided by the estimated population size.

Table 16-36
Number of Students Assessed in Mathematics Main Sample
Age 9/grade 4

	Grade			
	< 4	= 4	> 4	Total
AGE < 9				
Unweighted N	0	39	0	39
Estimated population size	0	17395	0	17395
Standard error	0	2980	0	2980
Coefficient of variation*	0.00	17.13	0.00	17.13
AGE = 9				
Unweighted N	2221	4359	17	6597
Estimated population size	1435596	1984715	11888	3432199
Standard error	13860	29417	3278	36417
Coefficient of variation*	0.97	1.48	27.57	1.06
AGE > 9				
Unweighted N	0	2778	0	2778
Estimated population size	0	1473087	0	1473087
Standard error	0	26989	0	26989
Coefficient of variation*	0.00	1.83	0.00	1.83
AGE TOTAL				
Unweighted N	2221	7176	17	9414
Estimated population size	1435596	3475197	11888	4922681
Standard error	13860	14376	3278	25474
Coefficient of variation*	0.97	0.41	27.57	0.52

---

* The coefficient of variation of the estimated population is defined as 100 times its standard error divided by the estimated population size.

Table 16-37
Number of Students Assessed in Mathematics Estimation Sample
Age 9/grade 4

	Grade < 4	Grade = 4	Grade > 4	Total
AGE < 9				
Unweighted N	0	5	0	5
Estimated population size	0	17327	0	17327
Standard error	0	7273	0	7273
Coefficient of variation*	0.00	41.97	0.00	41.97
AGE = 9				
Unweighted N	487	926	5	1418
Estimated population size	1421777	1962765	15268	3399810
Standard error	22429	54584	7547	59814
Coefficient of variation*	1.58	2.78	49.43	1.76
AGE > 9				
Unweighted N	0	631	0	631
Estimated population size	0	1470528	0	1470528
Standard error	0	57620	0	57620
Coefficient of variation*	0.00	3.92	0.00	3.92
AGE TOTAL				
Unweighted N	487	1562	5	2054
Estimated population size	1421777	3450619	15268	4887664
Standard error	22429	31149	7547	47475
Coefficient of variation*	1.58	0.90	49.43	0.97

---

* The coefficient of variation of the estimated population is defined as 100 times its standard error divided by the estimated population size.

Table 16-38

Number of Students Assessed in Mathematics Calculator Bridge Sample

Age 9/grade 4

	Grade			
	< 4	= 4	> 4	Total
AGE < 9				
Unweighted N	0	7	0	7
Estimated population size	0	11979	0	11979
Standard error	0	4553	0	4553
Coefficient of variation*	0.00	38.01	0.00	38.01
AGE = 9				
Unweighted N	567	1025	2	1594
Estimated population size	1453051	1999806	4924	3457780
Standard error	21191	46502	3291	50634
Coefficient of variation*	1.46	2.33	66.84	1.46
AGE > 9				
Unweighted N	0	635	0	635
Estimated population size	0	1462124	0	1462124
Standard error	0	48318	0	48318
Coefficient of variation*	0.00	3.30	0.00	3.30
AGE TOTAL				
Unweighted N	567	1667	2	2236
Estimated population size	1453051	3473909	4924	4931884
Standard error	21191	24699	3291	40603
Coefficient of variation*	1.46	0.71	66.84	0.82

---

* The coefficient of variation of the estimated population is defined as 100 times its standard error divided by the estimated population size.

Table 16-39
Number of Students Assessed in Writing Main Sample
Age 9/grade 4

	Grade			
	< 4	= 4	> 4	Total
AGE < 9				
Unweighted N	0	34	0	34
Estimated population size	0	16052	0	16052
Standard error	0	3904	0	3904
Coefficient of variation*	0.00	24.32	0.00	24.32
AGE = 9				
Unweighted N	2354	4366	32	6752
Estimated population size	1424121	2021789	18854	3464764
Standard error	14164	27332	3809	32169
Coefficient of variation*	0.99	1.35	20.20	0.93
AGE > 9				
Unweighted N	0	2766	0	2766
Estimated population size	0	1428402	0	1428402
Standard error	0	28803	0	28803
Coefficient of variation*	0.00	2.02	0.00	2.02
AGE TOTAL				
Unweighted N	2354	7166	32	9552
Estimated population size	1424121	3466243	18854	4909219
Standard error	14164	15272	3809	24470
Coefficient of variation*	0.99	0.44	20.20	0.50

---

* The coefficient of variation of the estimated population is defined as 100 times its standard error divided by the estimated population size.

Table 16-40
Number of Students Assessed in Reading Main Sample
Age 13/grade 8

	Grade			
	< 8	= 8	> 8	Total
AGE < 13				
Unweighted N	0	97	0	97
Estimated population size	0	26493	0	26493
Standard error	0	3527	0	3527
Coefficient of variation*	0.00	13.31	0.00	13.31
AGE = 13				
Unweighted N	3728	6821	30	10579
Estimated population size	1434572	1809193	31660	3275425
Standard error	19632	24417	14636	24714
Coefficient of variation*	1.37	1.35	46.23	0.75
AGE > 13				
Unweighted N	0	4266	0	4266
Estimated population size	0	1350391	0	1350391
Standard error	0	25890	0	25890
Coefficient of variation*	0.00	1.92	0.00	1.92
AGE TOTAL				
Unweighted N	3728	11184	30	14942
Estimated population size	1434572	3186077	31660	4652309
Standard error	19632	11027	14636	19588
Coefficient of variation*	1.37	0.35	46.23	0.42

---

* The coefficient of variation of the estimated population is defined as 100 times its standard error divided by the estimated population size.

Table 16-41
Number of Students Assessed in Mathematics Main Sample
Age 13/grade 8

	Grade			
	< 8	= 8	> 8	Total
AGE < 13				
Unweighted N	0	73	0	73
Estimated population size	0	26484	0	26484
Standard error	0	3497	0	3497
Coefficient of variation*	0.00	13.20	0.00	13.20
AGE = 13				
Unweighted N	2606	4662	22	7290
Estimated population size	1444947	1803754	23834	3272535
Standard error	14622	32021	11801	31215
Coefficient of variation*	1.01	1.78	49.51	0.95
AGE > 13				
Unweighted N	0	2928	0	2928
Estimated population size	0	1350660	0	1350660
Standard error	0	35019	0	35019
Coefficient of variation*	0.00	2.59	0.00	2.59
AGE TOTAL				
Unweighted N	2606	7663	22	10291
Estimated population size	1444947	3180899	23834	4649680
Standard error	14622	13152	11801	21954
Coefficient of variation*	1.01	0.41	49.51	0.47

---

* The coefficient of variation of the estimated population is defined as 100 times its standard error divided by the estimated population size.

Table 16-42
Number of Students Assessed in Mathematics Estimation Sample
Age 13/grade 8

	Grade			
	< 8	= 8	> 8	Total
AGE < 13				
Unweighted N	0	19	0	19
Estimated population size	0	35237	0	35237
Standard error	0	8966	0	8966
Coefficient of variation*	0.00	25.44	0.00	25.44
AGE = 13				
Unweighted N	642	1017	5	1664
Estimated population size	1474293	1764283	11722	3250298
Standard error	18231	40528	4818	42482
Coefficient of variation*	1.24	2.30	41.10	1.31
AGE > 13				
Unweighted N	0	733	0	733
Estimated population size	0	1402394	0	1402394
Standard error	0	45364	0	45364
Coefficient of variation*	0.00	3.23	0.00	3.23
AGE TOTAL				
Unweighted N	642	1769	5	2416
Estimated population size	1474293	3201913	11722	4687928
Standard error	18231	18167	4818	32519
Coefficient of variation*	1.24	0.57	41.10	0.69

---

* The coefficient of variation of the estimated population is defined as 100 times its standard error divided by the estimated population size.

Table 16-43
Number of Students Assessed in Writing Main Sample
Age 13/grade 8

	Grade			
	< 8	= 8	> 8	Total
AGE < 13				
Unweighted N	0	110	0	110
Estimated population size	0	29951	0	29951
Standard error	0	3812	0	3812
Coefficient of variation*	0.00	12.73	0.00	12.73
AGE = 13				
Unweighted N	3812	6803	18	10633
Estimated population size	1466581	1807922	9775	3284278
Standard error	10571	25971	4709	26391
Coefficient of variation*	0.72	1.44	48.17	0.80
AGE > 13				
Unweighted N	0	4199	0	4199
Estimated population size	0	1348908	0	1348908
Standard error	0	26864	0	26864
Coefficient of variation*	0.00	1.99	0.00	1.99
AGE TOTAL				
Unweighted N	3812	11112	18	14942
Estimated population size	1466581	3186782	9775	4663137
Standard error	10571	11351	4709	19854
Coefficient of variation*	0.72	0.36	48.17	0.43

---

* The coefficient of variation of the estimated population is defined as 100 times its standard error divided by the estimated population size.

Table 16-44
Number of Students Assessed in Reading Main Sample
Age 17/grade 12

	Grade			
	< 12	= 12	> 12	Total
AGE < 17				
Unweighted N	0	156	0	156
Estimated population size	0	28968	0	28968
Standard error	0	2712	0	2712
Coefficient of variation*	0.00	9.36	0.00	9.36
AGE = 17				
Unweighted N	3959	7722	0	11681
Estimated population size	1071384	1680447	0	2751832
Standard error	5287	14261	0	15524
Coefficient of variation*	0.49	0.85	0.00	0.56
AGE > 17				
Unweighted N	0	3477	0	3477
Estimated population size	0	825018	0	825018
Standard error	0	30406	0	30406
Coefficient of variation*	0.00	3.69	0.00	3.69
AGE TOTAL				
Unweighted N	3959	11355	0	15314
Estimated population size	1071384	2534434	0	3605818
Standard error	5287	21529	0	21432
Coefficient of variation*	0.49	0.85	0.00	0.59

---

* The coefficient of variation of the estimated population is defined as 100 times its standard error divided by the estimated population size.

Table 16-45
Number of Students Assessed in Mathematics Main Sample
Age 17/grade 12

	Grade			
	< 12	= 12	> 12	Total
AGE < 17				
Unweighted N	0	104	0	104
Estimated population size	0	30326	0	30326
Standard error	0	4100	0	4100
Coefficient of variation*	0.00	13.52	0.00	13.52
AGE = 17				
Unweighted N	2526	4802	0	7328
Estimated population size	1074356	1679032	0	2753387
Standard error	5906	17045	0	18913
Coefficient of variation*	0.55	1.02	0.00	0.69
AGE > 17				
Unweighted N	0	2067	0	2067
Estimated population size	0	812813	0	812813
Standard error	0	37327	0	37327
Coefficient of variation*	0.00	4.59	0.00	4.59
AGE TOTAL				
Unweighted N	2526	6973	0	9499
Estimated population size	1074356	2522170	0	3596526
Standard error	5906	27974	0	28484
Coefficient of variation*	0.55	1.11	0.00	0.79

---

* The coefficient of variation of the estimated population is defined as 100 times its standard error divided by the estimated population size.

Table 16-46
Number of Students Assessed in Mathematics Estimation Sample
Age 17/grade 12

	Grade			
	< 12	= 12	> 12	Total
AGE < 17				
Unweighted N	0	18	0	18
Estimated population size	0	27446	0	27446
Standard error	0	9984	0	9984
Coefficient of variation*	0.00	36.38	0.00	36.38
AGE = 17				
Unweighted N	548	1027	0	1575
Estimated population size	1070175	1664599	0	2734774
Standard error	10168	22091	0	26416
Coefficient of variation*	0.95	1.33	0.00	0.97
AGE > 17				
Unweighted N	0	481	0	481
Estimated population size	0	847904	0	847904
Standard error	0	54537	0	54537
Coefficient of variation*	0.00	6.43	0.00	6.43
AGE TOTAL				
Unweighted N	548	1526	0	2074
Estimated population size	1070175	2539949	0	3610124
Standard error	10168	48432	0	49001
Coefficient of variation*	0.95	1.91	0.00	1.36

---

* The coefficient of variation of the estimated population is defined as 100 times its standard error divided by the estimated population size.

Table 16-47
Number of Students Assessed in Writing Main Sample
Age 17/grade 12

	Grade			
	< 12	= 12	> 12	Total
AGE < 17				
Unweighted N	0	155	0	155
Estimated population size	0	27919	0	27919
Standard error	0	2645	0	2645
Coefficient of variation*	0.00	9.47	0.00	9.47
AGE = 17				
Unweighted N	4137	7927	0	12064
Estimated population size	1070648	1696248	0	2766896
Standard error	6097	12742	0	14337
Coefficient of variation*	0.57	0.75	0.00	0.52
AGE > 17				
Unweighted N	0	3450	0	3450
Estimated population size	0	820458	0	820458
Standard error	0	32083	0	32083
Coefficient of variation*	0.00	3.91	0.00	3.91
AGE TOTAL				
Unweighted N	4137	11532	0	15669
Estimated population size	1070648	2544625	0	3615273
Standard error	6097	24697	0	23933
Coefficient of variation*	0.57	0.97	0.00	0.66

---

* The coefficient of variation of the estimated population is defined as 100 times its standard error divided by the estimated population size.

Table 16-48
Number of Students Assessed in Reading and Writing Long-term Trend Sample
Age 9/grade 4

	Grade			
	< 4	= 4	> 4	Total
AGE < 9				
Unweighted N	0	15	0	15
Estimated population size	0	8973	0	8973
Standard error	0	2162	0	2162
Coefficient of variation*	0.00	24.09	0.00	24.09
AGE = 9				
Unweighted N	1879	3058	7	4944
Estimated population size	1475673	1962493	5361	3443526
Standard error	10597	39483	2736	42114
Coefficient of variation*	0.72	2.01	51.04	1.22
AGE > 9				
Unweighted N	0	2103	0	2103
Estimated population size	0	1546571	0	1546571
Standard error	0	36472	0	36472
Coefficient of variation*	0.00	2.36	0.00	2.36
AGE TOTAL				
Unweighted N	1879	5176	7	7062
Estimated population size	1475673	3518036	5361	4999070
Standard error	10597	16282	2736	21897
Coefficient of variation*	0.72	0.46	51.04	0.44

---

* The coefficient of variation of the estimated population is defined as 100 times its standard error divided by the estimated population size.

Table 16-49
Number of Students Assessed in Reading and Writing Long-term Trend Sample
Age 13/grade 8

	Grade			
	< 8	= 8	> 8	Total
AGE < 13				
Unweighted N	0	35	0	35
Estimated population size	0	30140	0	30140
Standard error	0	4896	0	4896
Coefficient of variation*	0.00	16.24	0.00	16.24
AGE = 13				
Unweighted N	1393	2554	18	3965
Estimated population size	1451228	1868622	43731	3363581
Standard error	19150	35953	20421	37303
Coefficient of variation*	1.32	1.92	46.70	1.11
AGE > 13				
Unweighted N	0	1514	0	1514
Estimated population size	0	1319124	0	1319124
Standard error	0	40759	0	40759
Coefficient of variation*	0.00	3.09	0.00	3.09
AGE TOTAL				
Unweighted N	1393	4103	18	5514
Estimated population size	1451228	3217887	43731	4712846
Standard error	19150	16819	20421	23625
Coefficient of variation*	1.32	0.52	46.70	0.50

---

* The coefficient of variation of the estimated population is defined as 100 times its standard error divided by the estimated population size.

Table 16-50
Number of Students Assessed in Reading and Writing Long-term Trend Sample
Age 17/grade 11

	Grade			
	< 12	= 12	> 12	Total
AGE < 17				
Unweighted N	0	331	0	331
Estimated population size	0	288465	0	288465
Standard error	0	24588	0	24588
Coefficient of variation*	0.00	8.52	0.00	8.52
AGE = 17				
Unweighted N	944	3217	286	4447
Estimated population size	852041	1918270	231597	3001908
Standard error	19550	5038	17325	10194
Coefficient of variation*	2.29	0.26	7.48	0.34
AGE > 17				
Unweighted N	0	791	0	791
Estimated population size	0	773753	0	773753
Standard error	0	27868	0	27868
Coefficient of variation*	0.00	3.60	0.00	3.60
AGE TOTAL				
Unweighted N	944	4339	286	5569
Estimated population size	852041	2980488	231597	4064126
Standard error	19550	13649	17325	17875
Coefficient of variation*	2.29	0.46	7.48	0.44

---

* The coefficient of variation of the estimated population is defined as 100 times its standard error divided by the estimated population size.

Table 16-51
Number of Students Assessed in Mathematics and Science Long-term Trend Samples
Age Only

	Grade			
	< 4	= 4	> 4	Total
AGE 9				
Unweighted N	2748	4569	18	7335
Estimated population size	1264839	2089132	8732	3362703
Standard error	40528	45890	2624	18824
Coefficient of variation*	3.20	2.20	30.05	0.56

	Grade			
	< 8	= 8	> 8	Total
AGE 13				
Unweighted N	2088	3813	8	5909
Estimated population size	1188442	1985195	7869	3181507
Standard error	34567	34556	3731	19787
Coefficient of variation*	2.91	1.74	47.41	0.62

	Grade			
	< 11	= 11	> 11	Total
AGE 17				
Unweighted N	943	3154	262	4359
Estimated population size	700891	2057599	178266	2936756
Standard error	31763	33052	13160	14852
Coefficient of variation*	4.53	1.61	7.38	0.51

---

* The coefficient of variation of the estimated population is defined as 100 times its standard error divided by the estimated population size.

Table 16-52
Weighted Percentage of Students in Reading Main Sample by Type of Eligibility and Subgroup Classification*
Age 9/grade 4

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Sex				
Male	50.93	51.06	46.88	52.70
Female	49.07	48.94	53.12	47.30
Race/Ethnicity				
White	71.27	70.58	72.38	70.33
Black	14.45	15.83	14.10	15.56
Hispanic	10.08	9.41	9.15	9.99
Other	4.20	4.18	4.37	4.12
Region				
Northeast	21.68	21.36	24.38	20.34
Southeast	24.07	23.40	21.51	24.65
Central	26.93	27.27	26.17	27.48
West	27.33	27.97	27.95	27.52
Parents' Education				
Less than high school	3.54	4.11	2.84	4.22
High school	12.58	12.40	11.98	12.70
Greater than high school	7.14	8.61	8.58	7.58
Graduated college	37.75	38.96	40.99	37.26
Unknown	38.60	35.70	35.43	37.86
Size and Type of Community				
Rural	12.36	12.26	11.75	12.53
Disadvantaged urban	8.69	8.95	8.25	9.05
Advantaged urban	10.29	9.76	11.53	9.40
Big city	9.06	9.17	10.46	8.56
Fringe	15.34	15.65	16.90	14.91
Medium city	13.99	14.46	13.10	14.69
Small places	30.28	29.76	28.01	30.85
Estimated Total Population	3476135	3464301	2028390	4912046

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-53
Weighted Percentage of Students in Mathematics Main Sample by Type of Eligibility and Subgroup Classification*
Age 9/grade 4

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Sex				
Male	49.81	50.21	45.57	51.80
Female	50.19	49.79	54.43	48.20
Race/Ethnicity				
White	70.74	70.39	71.71	70.11
Black	14.82	15.87	14.57	15.67
Hispanic	10.04	9.48	9.01	10.06
Other	4.39	4.26	4.71	4.17
Region				
Northeast	21.84	20.78	25.09	19.78
Southeast	23.97	24.26	22.69	24.69
Central	26.46	26.86	24.62	27.48
West	27.74	28.10	27.59	28.06
Parents' Education				
Less than high school	3.82	4.09	2.96	4.35
High school	11.32	12.34	11.50	11.97
Greater than high school	6.11	7.33	6.86	6.67
Graduated college	39.01	40.90	43.22	38.65
Unknown	39.66	35.23	35.43	38.24
Size and Type of Community				
Rural	11.67	12.23	10.96	12.35
Disadvantaged urban	9.26	9.37	8.17	9.78
Advantaged urban	12.61	12.02	14.21	11.55
Big city	8.18	7.95	9.50	7.49
Fringe	13.90	13.72	14.99	13.33
Medium city	15.45	15.47	14.21	15.96
Small places	28.93	29.25	27.95	29.55
Estimated Total Population	3432199	3475197	1984715	4922681

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-54
Weighted Percentage of Students in Mathematics Estimation Sample by Type of Eligibility and Subgroup Classification*
Age 9/grade 4

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Sex				
Male	48.18	48.45	44.97	49.66
Female	51.82	51.55	55.03	50.34
Race/Ethnicity				
White	71.10	70.37	72.19	70.15
Black	15.08	16.00	15.06	15.74
Hispanic	9.92	9.49	8.67	10.12
Other	3.90	4.14	4.08	4.00
Region				
Northeast	21.17	20.47	25.28	19.03
Southeast	25.23	25.85	21.83	27.03
Central	27.96	27.78	26.25	28.52
West	25.64	25.89	26.64	25.42
Parents' Education				
Less than high school	3.75	4.10	3.48	4.11
High school	13.25	12.92	12.94	13.14
Greater than high school	5.94	8.53	8.41	6.78
Graduated college	40.24	40.56	42.53	39.54
Unknown	36.68	33.59	32.39	36.22
Size and Type of Community				
Rural	7.98	8.17	6.82	8.58
Disadvantaged urban	8.64	8.59	7.51	9.05
Advantaged urban	11.04	9.74	12.96	9.35
Big city	6.14	6.58	7.00	6.11
Fringe	13.20	13.55	14.98	12.73
Medium city	18.46	19.22	18.52	18.97
Small places	34.55	34.15	32.21	35.20
Estimated Total Population	3399810	3450619	1962765	4887664

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-55

Weighted Percentage of Students in Mathematics Calculator Sample by Type of Eligibility and Subgroup Classification*

Age 9/grade 4

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Sex				
Male	49.81	50.63	46.43	51.76
Female	50.19	49.37	53.57	48.24
Race/Ethnicity				
White	70.23	70.59	70.76	70.27
Black	15.46	16.00	15.87	15.67
Hispanic	9.91	9.29	8.64	9.98
Other	4.40	4.12	4.72	4.08
Region				
Northeast	23.33	22.77	26.72	21.56
Southeast	24.35	24.70	22.80	25.22
Central	26.04	25.74	23.82	26.73
West	26.29	26.78	26.67	26.48
Parents' Education				
Less than high school	4.49	5.33	4.61	5.03
High school	12.78	12.76	12.38	12.93
Greater than high school	5.78	6.05	5.53	6.07
Graduated college	38.61	40.10	42.20	38.21
Unknown	38.34	35.59	35.27	37.64
Size and Type of Community				
Rural	10.48	10.77	8.27	11.58
Disadvantaged urban	10.75	10.09	9.66	10.73
Advantaged urban	13.15	13.68	16.26	12.26
Big city	7.84	7.42	9.17	7.01
Fringe	12.69	11.54	12.68	11.88
Medium city	15.93	17.10	15.86	16.78
Small places	29.16	29.40	28.10	29.76
Estimated Total Population	3457780	3473909	1999806	4931884

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-56
Weighted Percentage of Students in Writing Main Sample by Type of Eligibility and Subgroup Classification*
Age 9/grade 4

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Sex				
Male	49.27	49.26	44.56	51.20
Female	50.73	50.74	55.44	48.80
Race/Ethnicity				
White	70.95	70.52	71.71	70.33
Black	14.98	15.81	15.02	15.55
Hispanic	9.74	9.41	8.72	9.93
Other	4.32	4.26	4.55	4.19
Region				
Northeast	21.38	21.01	24.10	20.00
Southeast	23.80	23.33	21.78	24.31
Central	27.26	27.60	26.15	27.96
West	27.56	28.05	27.97	27.74
Parents' Education				
Less than high school	3.94	4.02	3.30	4.26
High school	13.48	13.84	13.88	13.57
Greater than high school	6.91	9.18	8.65	7.79
Graduated college	38.89	40.54	42.35	38.63
Unknown	36.51	32.19	31.53	35.51
Size and Type of Community				
Rural	10.83	10.28	9.25	11.09
Disadvantaged urban	8.25	8.48	7.72	8.63
Advantaged urban	11.74	11.09	13.35	10.62
Big city	8.10	8.21	9.23	7.71
Fringe	14.61	15.41	16.11	14.56
Medium city	14.98	16.05	14.28	16.02
Small places	31.50	30.48	30.06	31.38
Estimated Total Population	3464764	3466243	2021789	4909219

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-57
Weighted Percentage of Students in Reading Main Sample by Type of Eligibility and Subgroup Classification*
Age 13/grade 8

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Sex				
Male	50.37	51.08	45.95	52.57
Female	49.63	48.92	54.05	47.43
Race/Ethnicity				
White	70.12	69.81	72.33	69.05
Black	15.37	15.58	13.62	16.19
Hispanic	10.16	10.20	9.31	10.51
Other	4.35	4.42	4.74	4.25
Region				
Northeast	22.63	21.81	25.53	20.94
Southeast	22.86	24.87	22.07	24.54
Central	25.88	25.28	24.32	26.07
West	28.63	28.04	28.08	28.44
Parents' Education				
Less than high school	7.12	7.81	5.67	8.16
High school	23.29	23.67	21.79	24.14
Greater than high school	17.39	19.12	19.45	17.77
Graduated college	42.90	41.58	46.40	40.63
Unknown	9.18	7.63	6.56	9.14
Size and Type of Community				
Rural	7.72	7.48	7.03	7.82
Disadvantaged urban	10.08	10.16	9.44	10.38
Advantaged urban	9.74	10.35	12.17	9.22
Big city	9.14	9.02	8.80	9.19
Fringe	14.56	14.32	16.94	13.48
Medium city	17.35	16.48	14.28	17.95
Small places	31.41	32.19	31.35	31.96
Estimated Total Population	3275425	3186077	1809193	4652309

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-58
Weighted Percentage of Students in Mathematics Main Sample by Type of Eligibility and Subgroup Classification*
Age 13/grade 8

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Sex				
Male	50.38	51.33	45.53	52.91
Female	49.62	48.67	54.47	47.09
Race/Ethnicity				
White	70.51	69.89	73.01	69.12
Black	15.21	15.55	13.22	16.21
Hispanic	10.06	10.10	9.18	10.43
Other	4.22	4.46	4.58	4.25
Region				
Northeast	22.37	21.76	25.54	20.72
Southeast	23.81	25.20	22.91	25.11
Central	26.04	25.00	24.05	26.10
West	27.78	28.04	27.49	28.07
Parents' Education				
Less than high school	7.03	7.61	5.92	7.85
High school	22.17	23.60	20.54	23.79
Greater than high school	18.35	18.10	19.65	17.67
Graduated college	41.57	41.67	46.23	39.83
Unknown	10.45	8.71	7.47	10.42
Size and Type of Community				
Rural	8.99	9.24	8.29	9.44
Disadvantaged urban	8.47	8.70	7.58	8.98
Advantaged urban	10.18	10.44	12.98	9.26
Big city	9.19	8.95	8.79	9.18
Fringe	13.99	13.87	15.87	13.18
Medium city	16.96	15.88	14.74	17.08
Small places	32.22	32.91	31.74	32.88
Estimated Total Population	3272535	3180899	1803754	4649680

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-59
Weighted Percentage of Students in Mathematics Estimation Sample by Type of Eligibility and Subgroup Classification*
Age 13/grade 8

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Sex				
Male	53.31	54.08	48.48	55.65
Female	46.69	45.92	51.52	44.35
Race/Ethnicity				
White	69.86	69.55	72.30	68.73
Black	15.16	15.62	13.30	16.18
Hispanic	10.67	10.52	9.72	10.92
Other	4.31	4.30	4.68	4.17
Region				
Northeast	23.59	23.42	26.42	22.41
Southeast	22.57	23.56	21.43	23.68
Central	26.77	26.28	24.59	27.25
West	27.07	26.74	27.56	26.66
Parents' Education				
Less than high school	7.32	8.00	5.62	8.42
High school	23.53	23.68	20.76	24.67
Greater than high school	15.48	17.39	17.59	15.99
Graduated college	43.09	43.09	48.15	41.19
Unknown	10.33	7.84	7.87	9.56
Size and Type of Community				
Rural	9.44	9.41	9.20	9.51
Disadvantaged urban	12.74	11.15	10.65	12.44
Advantaged urban	17.23	18.72	23.94	15.72
Big city	6.05	7.15	6.55	6.62
Fringe	11.93	12.12	13.97	11.29
Medium city	18.27	16.56	14.25	18.61
Small places	24.33	24.90	21.43	25.81
Estimated Total Population	3250298	3201913	1764283	4687928

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-60
Weighted Percentage of Students in Writing Main Sample by Type of Eligibility and Subgroup Classification*
Age 13/grade 8

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Sex				
Male	50.89	50.41	45.31	52.73
Female	49.11	49.59	54.69	47.27
Race/Ethnicity				
White	70.09	69.82	72.37	69.02
Black	15.26	15.55	13.39	16.18
Hispanic	10.29	10.20	9.43	10.57
Other	4.36	4.42	4.80	4.23
Region				
Northeast	23.07	22.44	26.48	21.32
Southeast	22.65	24.75	21.25	24.63
Central	25.94	25.17	23.93	26.19
West	28.34	27.64	28.34	27.86
Parents' Education				
Less than high school	7.03	7.63	5.83	7.91
High school	23.80	24.48	21.67	25.09
Greater than high school	18.13	19.62	20.94	18.06
Graduated college	41.38	40.20	44.39	39.40
Unknown	9.28	7.90	7.02	9.21
Size and Type of Community				
Rural	8.90	8.78	8.16	9.11
Disadvantaged urban	9.60	9.98	8.70	10.21
Advantaged urban	8.70	9.02	10.44	8.24
Big city	9.11	9.25	9.17	9.18
Fringe	15.01	15.14	17.89	13.99
Medium city	17.55	16.56	14.71	17.97
Small places	31.14	31.27	30.94	31.31
Estimated Total Population	3284278	3186782	1807922	4663137

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-61
Weighted Percentage of Students in Reading Main Sample by Type of Eligibility and Subgroup Classification*
Age 17/grade 12

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Sex				
Male	49.70	48.63	44.93	51.17
Female	50.30	51.37	55.07	48.83
Race/Ethnicity				
White	70.40	71.77	74.14	69.62
Black	15.69	14.91	13.82	16.01
Hispanic	9.56	8.78	7.75	9.85
Other	4.34	4.54	4.29	4.51
Region				
Northeast	22.75	23.99	25.94	22.14
Southeast	24.10	23.10	22.86	23.98
Central	26.27	25.98	25.81	26.28
West	26.88	26.93	25.40	27.61
Parents' Education				
Less than high school	7.71	7.61	5.64	8.60
High school	23.57	21.68	20.93	23.47
Greater than high school	26.15	26.66	28.29	25.52
Graduated college	39.88	41.50	43.38	39.38
Unknown	2.47	2.32	1.58	2.78
Size and Type of Community				
Rural	10.22	10.10	9.72	10.37
Disadvantaged urban	11.14	10.26	9.32	11.37
Advantaged urban	11.02	11.86	13.38	10.51
Big city	7.26	7.48	7.76	7.19
Fringe	11.81	12.35	13.34	11.48
Medium city	13.08	12.45	11.86	13.20
Small places	35.46	35.51	34.62	35.89
Estimated Total Population	2751832	2534434	1680447	3605818

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-62

Weighted Percentage of Students in Mathematics Main Sample by Type of Eligibility and Subgroup Classification*

Age 17/grade 12

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Sex				
Male	49.81	48.79	44.64	51.51
Female	50.19	51.21	55.36	48.49
Race/Ethnicity				
White	70.80	71.10	74.84	69.12
Black	15.64	14.73	13.82	15.85
Hispanic	9.26	9.46	7.11	10.40
Other	4.31	4.71	4.23	4.62
Region				
Northeast	22.51	24.26	25.96	22.13
Southeast	24.95	23.49	23.45	24.63
Central	25.47	25.38	25.34	25.47
West	27.07	26.87	25.26	27.78
Parents' Education				
Less than high school	7.00	6.45	4.82	7.64
High school	22.19	21.41	19.46	22.91
Greater than high school	25.97	26.19	27.52	25.39
Graduated college	41.73	42.93	45.59	40.77
Unknown	2.80	2.64	2.22	2.96
Size and Type of Community				
Rural	11.83	11.83	10.90	12.27
Disadvantaged urban	11.31	9.99	9.51	11.23
Advantaged urban	11.48	11.94	13.56	10.83
Big city	7.95	8.22	8.32	7.96
Fringe	11.05	11.86	12.55	10.92
Medium city	13.86	13.69	12.79	14.24
Small places	32.51	32.47	32.37	32.55
Estimated Total Population	2753387	2522170	1679032	3596526

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-63
Weighted Percentage of Students in Mathematics Estimation Sample by Type of Eligibility and Subgroup Classification*
Age 17/grade 12

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Sex				
Male	49.85	51.47	45.56	52.97
Female	50.15	48.53	54.44	47.03
Race/Ethnicity				
White	70.67	70.24	75.08	68.34
Black	15.42	15.03	13.15	16.19
Hispanic	9.77	10.05	7.75	10.90
Other	4.13	4.68	4.02	4.57
Region				
Northeast	21.72	23.96	24.82	21.86
Southeast	25.83	23.81	23.89	25.31
Central	25.60	25.67	26.05	25.44
West	26.85	26.56	25.25	27.39
Parents' Education				
Less than high school	7.68	7.34	5.28	8.55
High school	25.39	23.84	22.10	25.82
Greater than high school	26.75	26.32	27.89	25.93
Graduated college	37.37	39.28	42.76	36.22
Unknown	2.37	2.74	1.45	3.05
Size and Type of Community				
Rural	5.79	5.82	6.09	5.68
Disadvantaged urban	13.01	12.77	10.48	14.01
Advantaged urban	11.50	12.08	14.54	10.51
Big city	7.44	7.56	7.27	7.60
Fringe	10.82	13.30	14.37	10.92
Medium city	11.88	12.32	11.25	12.48
Small places	39.57	36.14	35.99	38.81
Estimated Total Population	2734774	2539949	1664599	3610124

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-64
Weighted Percentage of Students in Writing Main Sample by Type of Eligibility and Subgroup Classification*
Age 17/grade 12

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Sex				
Male	49.59	48.85	44.48	51.47
Female	50.41	51.15	55.52	48.53
Race/Ethnicity				
White	70.66	71.29	74.51	69.29
Black	15.57	15.05	13.57	16.14
Hispanic	9.54	9.17	7.76	10.12
Other	4.23	4.49	4.16	4.45
Region				
Northeast	22.73	24.56	25.63	22.66
Southeast	24.48	22.88	23.57	23.78
Central	25.89	26.05	25.90	26.00
West	26.90	26.51	24.90	27.56
Parents' Education				
Less than high school	7.95	7.75	6.17	8.65
High school	22.81	22.15	21.17	23.11
Greater than high school	25.45	25.90	26.70	25.17
Graduated college	40.87	41.71	44.10	39.95
Unknown	2.62	2.21	1.61	2.80
Size and Type of Community				
Rural	10.35	10.29	10.14	10.40
Disadvantaged urban	11.32	10.25	9.38	11.48
Advantaged urban	10.79	11.49	12.94	10.28
Big city	7.38	7.80	7.69	7.53
Fringe	12.41	12.81	14.27	11.82
Medium city	13.03	12.46	11.64	13.28
Small places	34.73	34.88	33.95	35.20
Estimated Total Population	2766896	2544625	1696248	3615273

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-65
Weighted Percentage of Students in Reading and Writing Long-term Trend Sample by Type of Eligibility and Subgroup Classification*
Age 9/grade 4

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Sex				
Male	50.83	50.26	45.25	52.62
Female	49.17	49.74	54.75	47.38
Race/Ethnicity				
White	70.52	70.34	71.54	70.00
Black	14.52	15.76	13.85	15.65
Hispanic	10.94	10.09	10.14	10.66
Other	4.02	3.80	4.46	3.69
Region				
Northeast	21.12	21.29	24.69	19.83
Southeast	23.41	23.71	21.02	24.56
Central	28.26	27.74	27.24	28.29
West	27.21	27.26	27.05	27.31
Parents' Education				
Less than high school	5.01	4.75	4.37	5.08
High school	15.36	16.36	14.67	16.34
Greater than high school	4.96	5.59	5.06	5.36
Graduated college	39.34	41.49	43.63	39.17
Unknown	34.45	31.47	32.00	33.31
Size and Type of Community				
Rural	10.65	10.68	9.18	11.25
Disadvantaged urban	9.38	9.16	8.56	9.54
Advantaged urban	9.60	9.54	10.66	9.14
Big city	10.08	11.01	12.40	9.82
Fringe	15.92	15.06	16.38	15.14
Medium city	16.32	16.08	15.65	16.41
Small places	28.06	28.46	27.17	28.69
Estimated Total Population	3443526	3518036	1962493	4999070

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-66
Weighted Percentage of Students in Reading and Writing Long-term Trend Sample by Type of Eligibility and Subgroup Classification*
Age 13/grade 8

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Sex				
Male	49.24	49.48	43.31	51.76
Female	50.76	50.52	56.69	48.24
Race/Ethnicity				
White	69.54	70.34	71.10	69.47
Black	15.62	15.46	14.01	16.15
Hispanic	10.66	10.19	9.98	10.61
Other	4.18	4.01	4.92	3.78
Region				
Northeast	22.18	20.95	25.40	20.06
Southeast	23.12	24.60	21.28	24.86
Central	25.96	25.02	23.88	26.14
West	28.74	29.43	29.43	28.94
Parents' Education				
Less than high school	5.95	6.95	4.76	7.11
High school	28.14	28.83	26.80	29.15
Greater than high school	11.13	12.38	13.65	10.98
Graduated college	45.27	43.87	48.12	43.19
Unknown	9.07	7.78	6.53	9.19
Size and Type of Community				
Rural	9.54	10.52	8.05	10.80
Disadvantaged urban	10.57	8.82	8.52	10.19
Advantaged urban	11.11	11.71	15.33	9.84
Big city	12.97	12.96	14.34	12.42
Fringe	14.85	14.00	15.07	14.18
Medium city	18.34	17.94	15.87	19.05
Small places	22.61	24.05	22.81	23.51
Estimated Total Population	3363581	3217887	1868622	4712846

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-67
Weighted Percentage of Students in Reading and Writing Long-term Trend Sample by Type of Eligibility and Subgroup Classification*
Age 17/grade 11

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Sex				
Male	52.06	53.13	48.80	54.38
Female	47.94	46.87	51.20	45.62
Race/Ethnicity				
White	71.25	71.24	77.14	68.47
Black	14.82	13.85	10.46	16.16
Hispanic	9.99	10.55	8.22	11.24
Other	3.94	4.35	4.17	4.13
Region				
Northeast	21.66	22.00	21.59	21.94
Southeast	24.82	23.56	21.67	25.39
Central	24.51	25.76	27.21	24.15
West	29.01	28.68	29.53	28.52
Parents' Education				
Less than high school	8.08	8.24	6.19	9.09
High school	28.32	26.72	26.05	28.22
Greater than high school	19.41	19.69	21.00	18.86
Graduated college	41.05	42.58	44.75	40.42
Unknown	2.94	2.65	1.85	3.24
Size and Type of Community				
Rural	10.55	10.03	10.59	10.14
Disadvantaged urban	11.78	12.01	9.18	13.17
Advantaged urban	9.74	9.42	9.96	9.40
Big city	3.97	4.26	4.11	4.12
Fringe	21.21	21.11	21.56	20.97
Medium city	12.18	12.52	12.92	12.08
Small places	30.58	30.65	31.68	30.11
Estimated Total Population	3001908	2980488	1918270	4064126

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-68
Weighted Percentage of Students in Mathematics and Science Long-term Trend Sample by Type of Eligibility and Subgroup Classification*
Age 9

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Sex				
Male	48.59	45.56	45.56	48.59
Female	51.41	54.44	54.44	51.41
Race/Ethnicity				
White	71.23	73.41	73.41	71.23
Black	14.72	14.10	14.10	14.72
Hispanic	10.52	8.80	8.80	10.52
Other	3.53	3.69	3.69	3.53
Region				
Northeast	20.91	24.23	24.23	20.91
Southeast	24.14	21.92	21.92	24.14
Central	28.20	27.46	27.46	28.20
West	26.75	26.39	26.39	26.75
Parents' Education				
Less than high school	4.14	3.74	3.74	4.14
High school	13.47	12.50	12.50	13.47
Greater than high school	7.81	8.55	8.55	7.81
Graduated college	41.34	43.57	43.57	41.34
Unknown	32.83	31.27	31.27	32.83
Size and Type of Community				
Rural	9.46	8.29	8.29	9.46
Disadvantaged urban	8.18	7.58	7.58	8.18
Advantaged urban	10.85	12.39	12.39	10.85
Big city	9.59	11.33	11.33	9.59
Fringe	16.99	17.50	17.50	16.99
Medium city	15.90	14.97	14.97	15.90
Small places	29.03	27.95	27.95	29.03
Estimated Total Population	3362703	2089132	2089132	3362703

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-69
Weighted Percentage of Students in Mathematics and Science Long-term Trend Sample by Type of Eligibility and Subgroup Classification*
Age 13

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Sex				
Male	49.86	44.87	44.87	49.86
Female	50.14	55.13	55.13	50.14
Race/Ethnicity				
White	71.39	73.35	73.35	71.39
Black	14.94	13.88	13.88	14.94
Hispanic	10.13	8.73	8.73	10.13
Other	3.54	4.04	4.04	3.54
Region				
Northeast	20.87	22.97	22.97	20.87
Southeast	23.88	23.24	23.24	23.88
Central	26.65	25.83	25.83	26.65
West	28.60	27.95	27.95	28.60
Parents' Education				
Less than high school	5.92	4.78	4.78	5.92
High school	23.08	21.54	21.54	23.08
Greater than high school	18.37	20.63	20.63	18.37
Graduated college	44.03	46.33	46.33	44.03
Unknown	8.33	6.48	6.48	8.33
Size and Type of Community				
Rural	9.21	7.73	7.73	9.21
Disadvantaged urban	9.72	8.39	8.39	9.72
Advantaged urban	10.24	12.33	12.33	10.24
Big city	13.69	14.58	14.58	13.69
Fringe	15.67	16.99	16.99	15.67
Medium city	18.09	17.23	17.23	18.09
Small places	23.38	22.75	22.75	23.38
Estimated Total Population	3181507	1985195	1985195	3181507

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-70

Weighted Percentage of Students in Mathematics and Science Long-term Trend Sample by Type of Eligibility and Subgroup Classification*

Age 17

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Sex				
Male	50.67	48.13	48.13	50.67
Female	49.33	51.87	51.87	49.33
Race/Ethnicity				
White	71.72	76.58	76.58	71.72
Black	14.64	11.83	11.83	14.64
Hispanic	10.08	8.36	8.36	10.08
Other	3.56	3.23	3.23	3.56
Region				
Northeast	21.55	21.98	21.98	21.55
Southeast	24.57	21.56	21.56	24.57
Central	24.96	27.42	27.42	24.96
West	28.92	29.05	29.05	28.92
Parents' Education				
Less than high school	8.12	6.08	6.08	8.12
High school	21.38	20.11	20.11	21.38
Greater than high school	25.32	26.42	26.42	25.32
Graduated college	42.51	45.46	45.46	42.51
Unknown	2.48	1.75	1.75	2.48
Size and Type of Community				
Rural	10.65	10.91	10.91	10.65
Disadvantaged urban	11.59	9.34	9.34	11.59
Advantaged urban	10.25	10.29	10.29	10.25
Big city	3.90	4.01	4.01	3.90
Fringe	21.98	22.90	22.90	21.98
Medium city	12.68	12.62	12.62	12.68
Small places	28.96	29.94	29.94	28.96
Estimated Total Population	2936756	2057599	2057599	2936756

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-71
Weighted Percentage of Excluded Students in Main Sample by Type of Eligibility and Subgroup Classification*
Age 9/grade 4

	Eligible by			
	Age	Grade	Age & Grade	Total
Sex				
Male	62.90	63.38	57.02	64.38
Female	37.10	36.62	42.98	35.62
Race/Ethnicity				
White	48.97	48.81	41.87	50.33
Black	19.61	16.57	15.08	18.78
Hispanic	26.00	28.65	35.04	25.69
Other	5.42	5.96	8.02	5.20
Region				
Northeast	19.40	19.53	23.67	18.60
Southeast	25.77	23.37	15.48	26.48
Central	18.79	17.27	12.96	19.10
West	36.05	39.83	47.89	35.81
Size and Type of Community				
Rural	9.00	10.45	9.19	9.80
Disadvantaged urban	16.51	13.66	17.95	14.56
Advantaged urban	5.52	5.44	6.91	5.19
Big city	8.48	10.09	10.04	9.09
Fringe	17.01	17.64	20.02	16.76
Medium city	18.51	20.60	16.58	20.11
Small places	24.97	22.12	19.30	24.48
Reason for Exclusion				
Disability	65.57	67.58	57.38	68.41
Limited English Proficiency	24.24	26.16	36.57	22.83
Both disability and LEP	4.36	3.25	1.60	4.29
Nonreader	2.60	1.48	2.51	1.97
Other	3.22	1.53	1.94	2.51
Estimated Total Population	240118	222383	78526	383975

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-72
Weighted Percentage of Excluded Students in Main Sample by Type of Eligibility and Subgroup Classification*
Age 13/grade 8

	Eligible by			
	Age	Grade	Age & Grade	Total
Sex				
Male	63.31	64.37	58.96	64.59
Female	36.69	35.63	41.04	35.41
Race/Ethnicity				
White	48.35	50.84	44.03	50.41
Black	21.84	17.90	16.29	20.57
Hispanic	24.50	25.27	30.80	23.91
Other	5.31	6.00	8.88	5.11
Region				
Northeast	24.09	19.64	24.28	21.63
Southeast	18.72	21.51	13.91	21.02
Central	23.71	25.43	20.56	25.16
West	33.47	33.42	41.25	32.19
Size and Type of Community				
Rural	4.46	5.30	2.75	5.19
Disadvantaged urban	16.70	16.43	20.44	15.95
Advantaged urban	7.30	5.96	9.91	6.15
Big city	8.19	7.51	6.81	8.04
Fringe	19.46	18.16	25.32	17.80
Medium city	19.84	17.76	16.72	19.21
Small places	24.06	28.88	18.05	27.67
Reason for Exclusion				
Disability	70.54	70.17	56.44	72.61
Limited English Proficiency	24.30	25.29	38.02	22.64
Both disability and LEP	3.59	2.89	3.70	3.19
Nonreader	0.53	0.53	0.49	0.54
Other	0.87	0.99	1.16	0.89
Estimated Total Population	219373	195894	57543	357725

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-73
Weighted Percentage of Excluded Students in Main Sample by Type of Eligibility and Subgroup Classification*
Age 17/grade 12

	Eligible by			
	Age	Grade	Age & Grade	Total
Sex				
Male	62.28	61.86	59.59	62.44
Female	37.72	38.14	40.41	37.56
Race/Ethnicity				
White	44.81	54.36	54.84	48.14
Black	20.52	20.40	19.04	20.66
Hispanic	26.16	16.45	18.32	22.45
Other	8.51	8.79	7.80	8.75
Region				
Northeast	22.05	23.01	25.21	22.09
Southeast	19.84	20.32	14.21	20.84
Central	20.91	22.92	26.16	21.19
West	37.20	33.75	34.41	35.88
Size and Type of Community				
Rural	8.11	10.41	8.57	9.18
Disadvantaged urban	18.61	15.24	13.82	17.60
Advantaged urban	6.00	7.24	10.98	5.93
Big city	6.48	5.14	4.07	6.15
Fringe	11.81	10.87	11.46	11.39
Medium city	18.15	17.04	19.38	17.44
Small places	30.84	34.05	31.72	32.30
Reason for Exclusion				
Disability	68.71	76.55	69.43	72.47
Limited English Proficiency	26.04	18.60	23.24	22.76
Both disability and LEP	3.27	3.21	3.35	3.23
Nonreader	0.52	0.49	0.83	0.46
Other	1.34	1.15	3.15	1.00
Estimated Total Population	139368	106409	29273	216504

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-74
Weighted Percentage of Excluded Students in Long-term Trend Samples by Type of Eligibility and Subgroup Classification*
Age 9/grade 4

	Eligible by			
	Age	Grade	Age & Grade	Total
Sex				
Male	65.09	63.03	62.12	64.47
Female	34.91	36.97	37.88	35.53
Race/Ethnicity				
White	50.31	46.09	39.48	49.99
Black	18.32	19.89	17.13	19.48
Hispanic	18.51	20.70	22.50	19.01
Other	12.85	13.32	20.89	11.52
Region				
Northeast	28.95	24.70	32.18	25.80
Southeast	25.57	24.97	12.94	27.75
Central	11.89	13.62	12.22	12.84
West	33.58	36.71	42.66	33.61
Size and Type of Community				
Rural	9.59	9.28	5.24	10.28
Disadvantaged urban	17.42	16.52	19.10	16.55
Advantaged urban	10.53	13.46	18.34	10.70
Big city	9.37	9.31	7.82	9.65
Fringe	16.61	17.39	23.53	15.69
Medium city	15.39	14.33	13.67	15.11
Small places	21.09	19.69	12.29	22.03
Reason for Exclusion				
Disability	64.11	66.54	52.04	67.96
Limited English Proficiency	22.90	21.54	34.65	19.75
Both disability and LEP	4.64	4.39	5.35	4.35
Nonreader	5.16	4.48	5.31	4.72
Other	3.19	3.06	2.64	3.23
Estimated Total Population	241801	233875	79329	396347

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-75
Weighted Percentage of Excluded Students in Long-term Trend Samples by Type of Eligibility and Subgroup Classification*
Age 13/grade 8

	Eligible by			
	Age	Grade	Age & Grade	Total
Sex				
Male	62.27	64.48	60.30	63.78
Female	37.73	35.52	39.70	36.22
Race/Ethnicity				
White	41.78	38.76	38.12	40.71
Black	21.71	18.77	13.24	21.44
Hispanic	24.61	29.17	35.37	25.40
Other	11.90	13.29	13.27	12.45
Region				
Northeast	22.29	16.06	24.42	18.56
Southeast	29.64	30.63	18.30	31.96
Central	14.60	13.25	13.79	13.99
West	33.47	40.07	43.49	35.49
Size and Type of Community				
Rural	9.06	13.06	6.14	11.69
Disadvantaged urban	22.46	20.50	20.82	21.65
Advantaged urban	8.72	4.74	9.72	6.40
Big city	10.64	8.80	9.65	9.79
Fringe	15.98	16.26	18.50	15.74
Medium city	18.32	20.45	23.22	18.71
Small places	14.83	16.19	11.95	16.02
Reason for Exclusion				
Disability	58.61	56.28	51.01	58.53
Limited English Proficiency	26.63	28.16	39.22	25.49
Both disability and LEP	4.91	4.25	1.81	5.04
Nonreader	3.46	3.66	1.97	3.80
Other	6.39	5.89	5.99	6.18
Estimated Total Population	204474	181850	52356	333967

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-76
Weighted Percentage of Excluded Students in Long-term Trend Samples by Type of Eligibility and Subgroup Classification*
Age 17/grade 11

	Eligible by			
	Age	Grade	Age & Grade	Total
Sex				
Male	67.24	64.55	60.83	66.47
Female	32.76	35.45	39.17	33.53
Race/Ethnicity				
White	48.14	50.88	53.56	48.59
Black	19.88	17.09	15.97	19.01
Hispanic	21.11	16.48	18.39	19.23
Other	10.87	15.55	12.09	13.17
Region				
Northeast	22.12	27.13	26.23	24.50
Southeast	24.55	16.42	12.49	21.81
Central	20.03	22.45	22.66	20.93
West	33.29	34.00	38.62	32.76
Size and Type of Community				
Rural	9.36	11.85	8.89	10.68
Disadvantaged urban	22.24	13.56	17.42	18.57
Advantaged urban	10.15	13.95	20.36	10.66
Big city	2.96	1.34	1.33	2.34
Fringe	21.16	28.17	25.30	23.99
Medium city	9.88	11.74	8.24	11.36
Small places	24.25	19.40	18.46	22.41
Reason for Exclusion				
Disability	69.54	72.16	69.06	70.97
Limited English Proficiency	20.91	19.15	19.85	20.32
Both disability and LEP	4.25	4.70	3.04	4.74
Nonreader	3.05	3.19	5.56	2.63
Other	2.25	0.73	2.49	1.30
Estimated Total Population	169064	163394	47886	292533

---

* NAEP reporting subgroups are described in Appendix B.

Table 16-77
Number of Students in NAEP Main Sample with Proficiency Scores by Type of Eligibility

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Age 9/Grade 4				
Reading	5966	6314	3864	8416
Mathematics	6597	7176	4359	9414
Mathematics Estimation	1418	1562	926	2054
Writing	6752	7166	4366	9552
Total	20733	22218	13515	29436
Age 13/Grade 8				
Reading	8941	9464	5760	12645
Mathematics	7290	7663	4662	10291
Mathematics Estimation	1664	1769	1017	2416
Writing	10633	11112	6803	14942
Total	28528	30008	18242	40294
Age 17/Grade 12				
Reading	10132	9856	6697	13291
Mathematics	7328	6973	4802	9499
Mathematics Estimation	1575	1526	1027	2074
Writing	12064	11532	7927	15669
Total	31099	29887	20453	40533

Table 16-78
Number of Students in NAEP Long-term Trend Samples with Proficiency Scores by Type of Eligibility

	Eligible by			
	Age	Grade	Age & Grade	Age or Grade
Reading and Writing Long-term Trend				
Age 9/Grade 4				
Reading	4944	5176	3058	7062
Writing	4128	4290	2540	5878
Age 13/Grade 8				
Reading	3965	4103	2554	5514
Writing	3965	4103	2554	5514
Age 17/Grade 12				
Reading	4447	4339	3217	5569
Writing	4447	4339	3217	5569
Mathematics and Science Long-term Trend				
Age 9/Grade 4	7335	4569	4569	7335
Age 13/Grade 8	5909	3813	3813	5909
Age 17/Grade 12	4359	3154	3154	4359
Total				
Age 9/Grade 4	12279	9745	7627	14397
Age 13/Grade 8	9874	7916	6367	11423
Age 17/Grade 12	8806	7493	6371	9928

Table 16-79

Weighted Means, Standard Deviations, and Percentiles of Reading Distributions for Reading Main Samples, by Subgroup, Grade 4

	MEAN	STD DEV	5TH	10TH	25TH	50TH	75TH	90TH	95TH
-- TOTAL --	217.9(1.0)	36.4(0.6)	154.1(1.3)	169.6(1.6)	194.6(1.2)	220.4(1.1)	243.3(1.1)	262.9(1.5)	273.7(1.7)
GENDER									
MALE	213.9(1.2)	37.0(0.8)	149.3(1.9)	164.7(2.3)	190.3(1.9)	216.4(1.1)	240.0(1.8)	260.0(1.3)	270.8(1.7)
FEMALE	222.0(1.0)	35.3(0.7)	160.5(3.9)	175.5(2.3)	200.0(1.2)	224.0(1.1)	246.3(1.7)	265.8(2.8)	276.7(2.1)
RACE/ETHNICITY									
WHITE	225.9(1.2)	33.3(0.6)	168.3(2.9)	181.5(2.3)	204.7(1.3)	227.5(1.0)	249.0(1.4)	267.4(1.2)	277.6(1.5)
BLACK	193.3(1.7)	34.7(1.2)	133.9(3.4)	148.0(3.8)	170.2(2.0)	194.7(2.2)	217.1(1.8)	236.9(2.9)	248.8(3.1)
HISPANIC	202.0(2.2)	37.2(1.7)	137.9(4.7)	152.2(5.5)	177.4(2.9)	203.6(3.4)	227.8(2.3)	249.1(2.9)	259.9(3.6)
ASIAN/PACIFIC AMERICAN	215.5(3.3)	32.7(1.5)	160.7(8.6)	171.4(14.6)	193.5(6.9)	216.6(5.7)	239.6(7.0)	257.3(9.6)	267.9(9.0)
AMER IND/ALASKAN NATV	207.5(4.7)	38.6(3.8)	134.5(29.0)	156.5(11.0)	185.0(7.5)	212.0(12.6)	233.2(4.7)	251.4(6.0)	263.3(6.1)
PARENTS' EDUCATION LEVEL									
LESS THAN H.S.	198.8(2.7)	35.0(2.2)	141.8(5.2)	154.3(5.5)	175.6(3.7)	200.3(2.5)	222.7(3.6)	243.3(5.4)	254.5(9.9)
GRADUATED H.S.	213.3(1.7)	34.3(1.3)	153.0(3.2)	167.2(2.5)	192.2(1.9)	216.5(1.5)	237.2(1.3)	254.9(2.0)	265.2(3.1)
SOME EDUC AFTER H.S.	224.0(2.2)	34.4(1.4)	163.9(8.2)	178.9(2.6)	202.8(3.0)	225.4(2.8)	248.0(3.6)	267.0(5.2)	277.4(3.8)
GRADUATED COLLEGE	226.7(1.4)	36.1(0.8)	162.7(2.7)	177.7(3.3)	203.4(2.1)	229.5(1.7)	252.4(1.6)	271.1(2.4)	281.2(2.6)
UNKNOWN	211.0(1.2)	35.0(0.8)	149.2(1.1)	164.4(2.0)	189.1(1.4)	213.6(1.2)	234.9(1.6)	254.2(2.6)	264.9(2.2)

Table 16-80

Weighted Means, Standard Deviations, and Percentiles of Reading Distributions
for Reading Main Samples, by Subgroup, Grade 8

	MEAN	STD DEV	5TH	10TH	25TH	50TH	75TH	90TH	95TH
-- TOTAL --	260.2(0.9)	36.0(0.3)	197.0(1.8)	212.6(1.2)	236.7(1.2)	262.3(0.9)	285.3(1.0)	304.8(1.2)	316.0(1.3)
GENDER									
MALE	254.0(1.1)	35.9(0.4)	191.2(1.9)	205.8(1.5)	230.4(1.3)	256.1(1.2)	279.4(1.3)	298.6(1.9)	309.7(1.2)
FEMALE	266.7(1.0)	34.9(0.5)	206.7(1.4)	221.2(1.7)	243.9(1.2)	268.6(1.2)	291.0(1.2)	309.9(2.2)	320.8(1.7)
RACE/ETHNICITY									
WHITE	267.6(1.2)	33.4(0.4)	209.3(2.3)	223.4(2.1)	246.4(1.6)	269.5(1.1)	290.7(1.1)	308.9(1.5)	319.2(1.8)
BLACK	238.0(1.6)	33.1(0.7)	181.2(3.9)	194.5(3.1)	216.3(1.9)	239.5(1.3)	261.2(2.5)	279.1(1.2)	289.3(1.8)
HISPANIC	241.6(1.4)	36.5(0.8)	179.9(2.0)	192.6(2.1)	217.4(2.7)	242.5(1.4)	267.0(1.9)	287.5(2.2)	300.9(2.2)
ASIAN/PACIFIC AMERICAN	270.5(3.1)	36.3(2.0)	209.1(8.6)	223.2(2.7)	246.5(3.8)	270.7(3.5)	295.0(3.5)	318.4(4.9)	328.0(9.1)
AMER IND/ALASKAN NATV	251.2(3.7)	34.3(3.4)	192.4(7.9)	208.4(4.9)	230.8(3.2)	252.2(5.0)	274.5(12.0)	292.9(2.8)	303.5(11.1)
PARENTS' EDUCATION LEVEL									
LESS THAN H.S.	242.7(1.5)	34.3(1.7)	182.5(2.5)	197.5(3.0)	221.6(2.2)	243.8(2.2)	266.4(2.3)	286.3(3.6)	296.6(5.1)
GRADUATED H.S.	250.9(1.4)	34.1(0.7)	191.4(1.6)	205.4(3.1)	228.4(1.8)	253.0(2.0)	275.0(1.3)	293.2(1.8)	303.6(2.5)
SOME EDUC AFTER H.S.	265.6(1.1)	32.2(0.7)	209.8(3.3)	223.5(1.9)	244.9(1.9)	267.4(1.3)	287.3(1.5)	305.1(2.5)	315.2(2.3)
GRADUATED COLLEGE	271.0(1.0)	34.2(0.5)	210.8(2.4)	226.0(1.0)	249.3(1.4)	272.9(1.0)	294.9(1.1)	313.2(0.9)	323.3(1.8)
UNKNOWN	237.6(2.0)	36.0(1.1)	177.0(4.7)	190.2(3.6)	212.0(3.0)	238.6(2.5)	263.4(2.2)	284.0(1.8)	294.3(3.8)

496

497

Table 16-81

Weighted Means, Standard Deviations, and Percentiles of Reading Distributions for Reading Main Samples, by Subgroup, Grade 12

	MEAN	STD DEV	5TH	10TH	25TH	50TH	75TH	90TH	95TH
-- TOTAL --	291.0(0.6)	33.3(0.4)	233.0(1.2)	246.8(0.7)	269.2(0.8)	292.9(0.8)	314.5(0.6)	332.4(1.0)	342.6(0.8)
GENDER									
MALE	285.8(0.7)	33.3(0.4)	227.7(1.4)	241.3(1.4)	263.8(1.0)	287.7(1.0)	309.6(1.0)	327.3(0.8)	337.1(0.9)
FEMALE	296.0(0.7)	32.7(0.5)	238.8(1.4)	252.6(1.0)	274.6(1.3)	297.8(1.4)	318.7(1.0)	336.3(1.3)	346.5(0.8)
RACE/ETHNICITY									
WHITE	297.0(0.6)	31.4(0.5)	242.2(1.2)	255.6(1.6)	276.9(0.9)	298.7(0.8)	318.8(0.8)	335.6(0.7)	345.5(0.8)
BLACK	271.8(1.5)	31.2(0.8)	220.0(4.2)	231.0(2.1)	250.6(1.5)	272.2(2.0)	293.7(2.0)	312.0(1.4)	322.6(1.5)
HISPANIC	277.2(2.4)	32.9(1.1)	220.9(3.6)	233.2(2.6)	255.9(2.3)	278.7(2.9)	299.7(3.0)	318.4(3.5)	328.4(4.4)
ASIAN/PACIFIC AMERICAN	290.6(3.2)	36.7(1.8)	224.2(4.5)	241.1(2.6)	268.4(5.5)	293.7(3.7)	315.8(4.1)	336.2(2.6)	346.7(5.5)
AMER IND/ALASKAN NATV	271.8(5.3)	39.4(4.8)	210.4(8.3)	223.2(19.9)	241.2(7.2)	273.9(22.1)	302.5(24.1)	326.8(9.1)	337.0(9.2)
PARENTS' EDUCATION LEVEL									
LESS THAN H.S.	273.5(1.5)	31.5(0.9)	221.0(2.7)	231.5(2.5)	251.7(2.3)	274.0(2.1)	296.6(2.2)	314.1(1.7)	323.2(1.5)
GRADUATED H.S.	281.4(0.8)	31.7(0.6)	227.8(1.4)	239.9(1.7)	260.1(1.2)	282.2(0.9)	303.7(0.9)	321.6(1.7)	331.6(1.3)
SOME EDUC AFTER H.S.	292.7(0.8)	31.1(0.6)	238.7(2.4)	251.4(2.0)	272.3(1.1)	294.4(1.3)	314.4(0.8)	330.8(1.6)	341.1(1.9)
GRADUATED COLLEGE	300.5(0.8)	31.7(0.5)	244.6(1.6)	258.7(1.4)	280.7(0.9)	302.4(1.0)	322.7(1.1)	339.2(1.0)	348.8(1.3)
UNKNOWN	256.9(2.8)	34.3(2.5)	197.1(10.0)	210.4(10.7)	233.1(5.1)	259.0(7.7)	281.6(2.9)	300.0(5.2)	309.6(6.0)

Table 16-82

Percentiles for Reading Long-term Trend, Grade 4

	1971	1975	1980	1984	1988	1990	1992
TOTAL SAMPLE							
MEAN	207.6 (1.0)	210.0 (0.7)	215.0 (1.0)	211.0 (0.9)	211.8 (1.1)	209.2 (1.2)	210.5 (0.9)
ST. DEV.	42.1 (0.4)	38.6 (0.3)	37.9 (0.4)	41.1 (0.4)	41.2 (1.0)	44.7 (0.8)	40.3 (0.6)
PERCENTILES							
5	134.8 (2.0)	143.2 (1.3)	148.5 (1.6)	140.5 (1.2)	141.9 (3.6)	134.8 (3.2)	140.7 (1.6)
10	151.6 (1.6)	159.2 (1.1)	165.1 (1.4)	156.7 (1.2)	156.7 (2.1)	150.1 (1.9)	156.0 (1.5)
25	180.0 (1.3)	185.2 (0.8)	191.1 (1.2)	183.7 (1.2)	184.3 (1.8)	178.7 (1.8)	183.1 (1.5)
50	209.3 (1.0)	211.9 (0.8)	217.2 (0.9)	212.6 (1.0)	213.7 (1.4)	210.3 (1.5)	213.6 (0.9)
75	236.7 (1.0)	236.5 (0.9)	241.3 (1.0)	239.6 (0.9)	240.1 (1.3)	240.3 (1.8)	239.3 (1.2)
90	260.5 (0.8)	258.1 (0.8)	261.7 (1.1)	262.8 (0.9)	263.0 (1.7)	265.7 (1.8)	259.9 (1.2)
95	274.1 (0.9)	270.6 (1.1)	273.3 (1.6)	276.5 (1.4)	277.5 (2.0)	280.4 (1.3)	272.1 (1.2)
MALE STUDENTS							
MEAN	201.2 (1.1)	204.3 (0.8)	210.0 (1.1)	207.5 (1.0)	207.5 (1.4)	204.0 (1.7)	205.9 (1.3)
ST. DEV.	42.1 (0.5)	39.0 (0.5)	38.7 (0.5)	42.3 (0.5)	42.7 (1.2)	45.1 (1.0)	41.3 (0.6)
PERCENTILES							
5	128.9 (2.0)	136.6 (1.1)	141.9 (2.3)	136.0 (1.1)	136.5 (2.9)	129.6 (5.8)	136.6 (2.0)
10	145.0 (1.7)	152.6 (1.3)	158.7 (1.4)	151.1 (1.5)	151.1 (2.4)	145.1 (1.9)	150.5 (2.1)
25	173.6 (1.4)	178.9 (1.0)	185.3 (1.4)	178.5 (1.1)	178.4 (1.8)	172.2 (2.8)	176.8 (1.7)
50	202.8 (1.2)	206.1 (0.9)	212.5 (1.2)	209.1 (1.3)	209.7 (1.8)	204.4 (2.2)	208.3 (1.6)
75	230.4 (1.1)	231.4 (1.0)	237.1 (1.1)	237.7 (1.2)	237.1 (1.9)	236.1 (1.9)	235.5 (1.6)
90	254.6 (1.2)	253.0 (1.1)	257.5 (0.8)	261.1 (1.1)	260.3 (2.0)	261.7 (2.6)	257.3 (1.0)
95	268.4 (1.5)	265.4 (1.4)	268.7 (1.1)	275.1 (1.1)	275.1 (2.3)	276.1 (5.6)	269.9 (2.8)
FEMALE STUDENTS							
MEAN	213.9 (1.0)	215.8 (0.8)	220.1 (1.1)	214.4 (0.9)	216.3 (1.3)	214.5 (1.2)	215.4 (0.9)
ST. DEV.	41.0 (0.6)	37.3 (0.4)	36.5 (0.5)	39.6 (0.5)	39.2 (1.2)	43.6 (1.3)	38.8 (0.9)
PERCENTILES							
5	142.9 (2.1)	151.3 (1.4)	157.1 (1.8)	146.4 (2.1)	149.3 (5.5)	140.6 (3.9)	147.3 (3.3)
10	159.5 (1.3)	167.1 (1.1)	172.5 (1.7)	162.9 (1.6)	164.3 (4.6)	156.8 (3.2)	163.8 (2.3)
25	186.7 (1.2)	192.0 (1.0)	197.2 (1.2)	188.7 (1.0)	190.6 (2.4)	185.7 (1.7)	190.0 (1.4)
50	215.6 (1.1)	217.2 (0.9)	221.7 (1.1)	215.7 (1.0)	217.5 (2.0)	215.9 (1.3)	218.5 (1.3)
75	242.4 (1.1)	241.1 (1.0)	245.2 (1.1)	241.6 (1.0)	242.6 (1.1)	244.4 (1.9)	242.5 (1.3)
90	265.0 (0.9)	262.3 (1.0)	265.5 (1.7)	264.4 (1.3)	265.3 (2.2)	269.4 (1.9)	262.3 (1.9)
95	278.6 (1.5)	274.8 (1.1)	277.0 (1.5)	277.8 (2.0)	279.1 (3.4)	284.1 (2.1)	273.7 (2.0)

Table 16-82 (continued)
Percentiles for Reading Long-term Trend, Grade 4

	1971	1975	1980	1984	1988	1990	1992
WHITE STUDENTS							
MEAN	214.0 (0.9)	216.6 (0.7)	221.3 (0.8)	218.2 (0.9)	217.7 (1.4)	217.0 (1.3)	217.9 (1.0)
ST. DEV.	39.4 (0.4)	36.1 (0.3)	35.2 (0.3)	38.8 (0.3)	39.3 (1.0)	42.9 (1.0)	37.5 (0.7)
PERCENTILES							
5	146.3 (2.4)	154.4 (1.2)	160.7 (1.5)	152.0 (1.3)	150.2 (3.4)	144.2 (3.2)	152.8 (1.6)
10	162.4 (1.3)	169.8 (1.1)	175.3 (1.0)	167.1 (1.0)	165.0 (3.9)	160.0 (1.5)	167.0 (1.8)
25	188.1 (1.2)	193.3 (0.7)	199.0 (0.9)	192.4 (1.0)	191.8 (2.4)	188.0 (2.8)	192.8 (1.3)
50	215.2 (0.9)	217.9 (0.7)	222.8 (0.8)	219.5 (1.0)	219.1 (1.2)	218.4 (2.1)	220.6 (1.3)
75	241.0 (0.9)	241.0 (0.9)	245.7 (0.9)	244.9 (0.9)	244.3 (1.8)	246.7 (2.3)	244.2 (1.2)
90	263.6 (0.8)	261.6 (1.0)	265.1 (1.1)	267.2 (1.3)	266.7 (2.2)	270.9 (2.1)	264.0 (1.0)
95	276.7 (0.9)	273.8 (1.3)	276.4 (1.2)	280.2 (1.3)	280.6 (2.6)	285.3 (2.6)	275.9 (3.1)
BLACK STUDENTS							
MEAN	170.1 (1.7)	181.2 (1.2)	189.3 (1.8)	185.7 (1.4)	188.5 (2.4)	181.8 (2.9)	184.5 (2.2)
ST. DEV.	38.3 (0.7)	35.8 (0.6)	37.6 (1.0)	38.9 (0.9)	39.4 (1.6)	41.7 (1.7)	39.8 (1.3)
PERCENTILES							
5	106.7 (2.5)	118.8 (2.3)	123.1 (4.1)	120.8 (2.2)	124.7 (6.3)	115.0 (4.7)	119.3 (6.1)
10	120.0 (2.0)	133.7 (2.8)	139.4 (4.0)	135.1 (2.8)	138.3 (3.4)	128.9 (3.9)	132.4 (3.6)
25	143.4 (2.0)	157.5 (2.3)	165.3 (1.9)	159.3 (1.8)	161.8 (3.0)	152.5 (3.2)	156.3 (4.4)
50	171.0 (2.1)	182.8 (1.2)	191.7 (2.1)	186.5 (1.5)	188.3 (4.0)	181.8 (3.1)	185.1 (2.5)
75	196.3 (1.8)	206.5 (1.2)	215.6 (1.9)	212.5 (1.6)	216.5 (2.9)	210.5 (2.4)	213.5 (2.6)
90	218.9 (1.6)	226.3 (1.5)	236.3 (1.9)	235.3 (2.5)	238.2 (3.8)	236.3 (2.7)	235.5 (2.9)
95	232.4 (1.7)	237.2 (2.0)	247.1 (1.8)	248.4 (2.0)	252.2 (4.6)	250.7 (6.9)	248.7 (2.5)
HISPANIC STUDENTS *							
MEAN	0.0 (0.0)	182.7 (2.2)	190.2 (2.3)	187.1 (3.1)	193.7 (3.5)	189.4 (2.3)	191.7 (3.1)
ST. DEV.	0.0 (0.0)	36.8 (1.3)	38.2 (1.2)	39.2 (1.5)	41.5 (2.8)	39.7 (1.6)	40.3 (1.8)
PERCENTILES							
5	0.0 (0.0)	120.3 (4.9)	123.4 (3.1)	120.3 (5.1)	121.8 (11.3)	125.4 (8.9)	124.8 (6.2)
10	0.0 (0.0)	133.4 (5.2)	138.4 (4.1)	134.7 (7.2)	140.3 (7.7)	139.0 (4.3)	138.7 (5.6)
25	0.0 (0.0)	157.4 (3.0)	164.3 (3.9)	160.7 (2.4)	164.9 (5.1)	160.8 (1.9)	162.5 (6.0)
50	0.0 (0.0)	184.2 (2.9)	192.0 (3.3)	189.2 (2.3)	196.0 (3.4)	189.3 (3.5)	192.7 (4.6)
75	0.0 (0.0)	209.4 (3.4)	217.6 (3.0)	215.4 (2.3)	222.0 (6.0)	218.9 (4.0)	222.0 (2.3)
90	0.0 (0.0)	228.6 (3.6)	237.8 (2.7)	236.1 (2.2)	246.7 (8.0)	239.3 (5.7)	244.7 (5.6)
95	0.0 (0.0)	240.3 (2.6)	249.9 (4.3)	247.1 (2.1)	258.6 (11.4)	253.2 (6.7)	255.4 (10.4)

* NO DATA WERE AVAILABLE FOR HISPANIC STUDENTS IN 1971

Table 16-83

Percentiles for Reading Long-term Trend, Grade 8

	1971	1975	1980	1984	1988	1990	1992
TOTAL SAMPLE							
MEAN	255.2 (0.9)	255.9 (0.8)	258.5 (0.9)	257.1 (0.6)	257.5 (1.0)	256.8 (0.8)	259.8 (1.2)
ST. DEV.	35.7 (0.4)	35.8 (0.3)	34.9 (0.4)	35.5 (0.3)	34.7 (0.5)	36.0 (0.6)	39.4 (0.8)
PERCENTILES							
5	192.8 (1.8)	193.5 (1.1)	199.1 (1.9)	196.7 (1.1)	199.5 (1.7)	195.7 (1.9)	190.9 (2.8)
10	207.8 (1.4)	208.7 (1.0)	212.8 (1.5)	210.2 (0.9)	212.9 (1.2)	209.8 (1.8)	207.9 (1.9)
25	232.3 (1.2)	232.9 (1.0)	235.3 (1.1)	233.9 (0.8)	234.2 (1.2)	233.2 (1.0)	234.7 (1.8)
50	257.0 (1.0)	257.7 (0.9)	259.6 (0.8)	258.2 (0.8)	257.9 (1.1)	257.3 (0.9)	261.6 (1.6)
75	279.9 (0.8)	280.6 (0.8)	282.8 (0.8)	281.6 (0.6)	281.4 (1.4)	281.5 (0.8)	287.0 (1.4)
90	299.6 (0.9)	300.5 (1.0)	302.3 (0.8)	301.7 (0.8)	301.6 (1.0)	302.0 (1.0)	309.2 (1.8)
95	310.8 (0.9)	311.8 (1.0)	313.9 (0.8)	313.7 (1.0)	313.7 (1.3)	314.4 (1.3)	321.9 (2.6)
MALE STUDENTS							
MEAN	249.6 (1.0)	249.6 (0.8)	254.3 (1.1)	252.7 (0.7)	251.8 (1.3)	250.5 (1.1)	254.1 (1.7)
ST. DEV.	35.9 (0.5)	35.7 (0.4)	35.0 (0.5)	35.8 (0.4)	35.3 (0.7)	36.0 (0.7)	40.4 (1.1)
PERCENTILES							
5	186.7 (1.6)	187.2 (1.1)	194.9 (1.9)	191.9 (1.0)	192.6 (2.5)	189.7 (2.2)	184.9 (3.7)
10	201.6 (1.6)	202.3 (1.5)	208.5 (1.5)	205.5 (1.2)	206.7 (1.8)	202.8 (1.4)	201.0 (2.8)
25	226.3 (1.2)	226.8 (1.1)	230.8 (1.2)	228.9 (1.1)	227.7 (2.1)	226.9 (1.9)	227.4 (2.1)
50	251.4 (0.8)	251.4 (0.9)	255.4 (1.1)	253.9 (0.9)	252.1 (2.1)	251.9 (1.3)	255.6 (2.4)
75	274.5 (0.8)	274.1 (0.8)	278.6 (1.2)	277.5 (1.0)	276.5 (2.0)	275.3 (1.2)	282.5 (1.5)
90	294.2 (1.0)	293.5 (1.0)	298.5 (1.2)	297.8 (1.0)	297.2 (1.5)	295.3 (1.2)	305.0 (3.4)
95	305.9 (1.3)	305.6 (1.7)	309.9 (0.9)	309.4 (1.2)	309.4 (2.8)	307.4 (3.2)	317.8 (3.1)
FEMALE STUDENTS							
MEAN	260.8 (0.9)	262.3 (0.9)	262.6 (0.9)	261.8 (0.7)	263.0 (1.0)	263.1 (1.1)	265.3 (1.2)
ST. DEV.	34.5 (0.4)	34.8 (0.4)	34.2 (0.4)	34.5 (0.3)	33.1 (0.6)	34.8 (0.7)	37.5 (0.8)
PERCENTILES							
5	200.9 (1.5)	202.1 (1.7)	204.2 (2.0)	203.0 (1.3)	207.3 (3.9)	205.3 (3.1)	199.3 (4.1)
10	215.2 (1.4)	215.9 (1.4)	218.0 (2.0)	216.8 (1.1)	221.0 (1.6)	217.9 (2.0)	216.8 (2.9)
25	238.5 (0.8)	239.8 (1.1)	240.0 (1.1)	239.1 (0.8)	240.0 (1.6)	240.0 (1.9)	241.5 (1.2)
50	262.4 (1.1)	264.2 (1.0)	263.4 (0.9)	262.7 (0.8)	263.0 (1.4)	263.0 (1.6)	266.6 (1.9)
75	285.0 (1.0)	286.6 (1.2)	286.3 (1.0)	285.4 (0.7)	285.8 (1.0)	286.6 (1.1)	290.8 (1.1)
90	303.8 (1.3)	305.4 (1.0)	305.6 (1.0)	305.5 (0.8)	305.2 (1.2)	308.1 (1.5)	312.8 (1.5)
95	314.6 (0.9)	316.1 (1.1)	317.3 (1.6)	317.5 (1.6)	317.7 (3.2)	319.4 (2.5)	324.5 (2.4)

Table 16-83 (continued)
Percentiles for Reading Long-term Trend, Grade 8

	1971	1975	1980	1984	1988	1990	1992
WHITE STUDENTS							
MEAN	260.9 (0.7)	262.1 (0.7)	264.4 (0.7)	262.5 (0.6)	261.3 (1.1)	262.3 (0.9)	266.4 (1.2)
ST. DEV.	32.9 (0.3)	32.9 (0.3)	32.7 (0.3)	33.8 (0.4)	33.9 (0.5)	34.5 (0.6)	36.6 (0.7)
PERCENTILES							
5	204.6 (1.2)	206.3 (1.0)	209.0 (1.2)	204.9 (0.9)	204.0 (1.4)	204.1 (2.2)	204.0 (2.7)
10	217.9 (0.9)	219.2 (0.7)	221.8 (1.2)	218.3 (0.8)	217.1 (2.1)	217.3 (1.7)	218.7 (2.2)
25	239.4 (0.9)	240.7 (0.8)	242.8 (0.8)	240.6 (0.8)	238.3 (1.0)	239.6 (1.7)	242.5 (1.4)
50	262.0 (0.8)	263.1 (1.0)	265.1 (0.6)	263.4 (0.7)	262.2 (1.1)	262.6 (1.4)	267.5 (2.0)
75	283.5 (0.9)	284.6 (0.8)	286.9 (0.7)	285.6 (0.7)	285.1 (0.9)	285.6 (1.2)	291.5 (1.1)
90	302.2 (0.7)	303.5 (0.9)	305.7 (0.8)	305.0 (0.8)	304.2 (1.5)	306.0 (2.4)	312.4 (1.9)
95	313.1 (1.1)	314.3 (0.9)	316.9 (0.8)	316.8 (1.3)	315.8 (1.1)	318.1 (2.7)	324.4 (2.0)
BLACK STUDENTS							
MEAN	222.4 (1.2)	225.7 (1.2)	232.8 (1.5)	236.3 (1.2)	242.9 (2.4)	241.5 (2.2)	237.6 (2.3)
ST. DEV.	33.5 (0.5)	34.9 (0.7)	32.7 (0.8)	34.1 (0.8)	32.1 (1.3)	35.3 (1.5)	39.8 (1.9)
PERCENTILES							
5	166.3 (1.5)	167.2 (2.5)	178.6 (2.4)	180.1 (2.0)	190.6 (3.4)	182.3 (5.3)	169.6 (10.1)
10	178.0 (2.2)	180.1 (2.5)	190.6 (3.3)	192.4 (1.9)	202.2 (3.3)	194.3 (7.3)	185.3 (3.3)
25	199.1 (1.9)	202.2 (1.3)	210.9 (1.8)	213.3 (2.6)	222.0 (2.4)	217.0 (3.2)	210.0 (3.0)
50	223.3 (1.4)	226.0 (1.7)	232.6 (1.3)	236.4 (1.3)	242.4 (2.7)	242.5 (4.0)	239.2 (2.3)
75	245.5 (1.4)	249.9 (1.5)	254.8 (1.9)	259.3 (1.1)	263.6 (4.5)	265.7 (2.5)	265.6 (2.8)
90	264.8 (1.3)	270.6 (1.2)	275.0 (1.7)	280.3 (1.9)	283.6 (4.7)	285.9 (4.9)	287.3 (3.1)
95	276.8 (2.3)	282.7 (2.3)	286.2 (1.5)	292.7 (1.6)	298.9 (2.2)	298.9 (3.0)	302.5 (4.9)
HISPANIC STUDENTS *							
MEAN	0.0 (0.0)	232.5 (3.0)	237.2 (2.0)	239.6 (2.0)	240.1 (3.5)	237.8 (2.3)	239.2 (3.5)
ST. DEV.	0.0 (0.0)	34.5 (1.0)	32.7 (0.8)	34.9 (1.2)	34.6 (2.4)	35.9 (1.3)	40.4 (2.4)
PERCENTILES							
5	0.0 (0.0)	173.7 (6.9)	182.6 (4.8)	180.8 (2.9)	181.4 (6.9)	178.0 (9.6)	165.0 (13.0)
10	0.0 (0.0)	186.7 (2.8)	194.9 (4.5)	193.3 (3.3)	194.6 (3.8)	191.3 (4.9)	183.8 (8.0)
25	0.0 (0.0)	207.8 (3.0)	214.8 (3.0)	216.1 (2.5)	218.9 (6.1)	214.1 (4.1)	213.0 (5.7)
50	0.0 (0.0)	233.5 (3.6)	237.5 (2.4)	240.4 (2.5)	240.3 (4.1)	238.6 (4.1)	242.0 (10.6)
75	0.0 (0.0)	256.7 (4.8)	259.3 (1.9)	263.5 (2.3)	262.0 (5.4)	262.2 (3.1)	267.0 (7.7)
90	0.0 (0.0)	277.2 (2.3)	279.2 (2.9)	284.2 (2.2)	284.0 (8.7)	283.8 (6.0)	288.7 (8.0)
95	0.0 (0.0)	289.1 (3.5)	290.5 (1.5)	295.9 (3.1)	297.3 (10.1)	295.9 (4.5)	303.1 (7.7)

* NO DATA WERE AVAILABLE FOR HISPANIC STUDENTS IN 1971

Table 16-84

Percentiles for Reading Long-term Trend, Grade 12

	1971	1975	1980	1984	1988	1990	1992
TOTAL SAMPLE							
MEAN	285.2 (1.2)	285.6 (0.8)	285.6 (0.8)	288.8 (0.8)	290.1 (1.0)	290.2 (1.1)	285.5 (1.2)
ST. DEV.	45.8 (0.5)	44.0 (0.6)	44.0 (0.6)	40.3 (0.3)	37.1 (0.7)	41.3 (0.7)	41.8 (0.6)
PERCENTILES							
5	206.1 (1.5)	209.3 (3.0)	209.3 (3.0)	219.9 (1.3)	226.1 (1.3)	220.0 (2.3)	213.0 (1.7)
10	225.3 (1.7)	228.4 (1.7)	228.4 (1.7)	236.0 (0.9)	241.5 (2.2)	236.9 (3.1)	230.6 (1.8)
25	255.9 (1.6)	257.8 (1.1)	257.8 (1.1)	262.5 (1.1)	265.7 (1.8)	263.5 (1.3)	258.7 (1.2)
50	287.7 (1.4)	287.9 (0.7)	287.9 (0.7)	290.3 (0.9)	291.1 (1.9)	291.1 (1.3)	287.5 (1.4)
75	316.7 (1.0)	315.7 (0.7)	315.7 (0.7)	316.8 (0.9)	316.0 (1.4)	318.6 (1.5)	314.6 (1.2)
90	341.7 (1.1)	340.0 (0.9)	340.0 (0.9)	339.6 (0.7)	336.9 (2.1)	342.7 (2.1)	337.5 (1.4)
95	356.5 (1.5)	354.3 (0.7)	354.3 (0.7)	352.6 (1.0)	348.7 (1.8)	356.0 (1.7)	350.9 (1.3)
MALE STUDENTS							
MEAN	278.9 (1.2)	279.7 (1.0)	279.7 (1.0)	283.9 (0.8)	286.0 (1.5)	284.0 (1.6)	281.8 (1.3)
ST. DEV.	46.3 (0.6)	45.1 (0.6)	45.1 (0.6)	40.9 (0.4)	37.5 (1.2)	42.6 (0.8)	42.7 (0.6)
PERCENTILES							
5	198.3 (1.6)	201.6 (1.4)	201.6 (1.4)	214.3 (1.5)	222.0 (2.3)	209.4 (3.2)	207.2 (1.9)
10	218.2 (2.0)	220.8 (2.0)	220.8 (2.0)	230.1 (1.0)	236.3 (3.7)	228.2 (3.4)	225.4 (2.2)
25	249.1 (1.4)	250.9 (1.1)	250.9 (1.1)	257.0 (1.3)	261.6 (1.8)	257.3 (1.9)	254.4 (1.5)
50	281.6 (1.4)	282.0 (1.3)	282.0 (1.3)	285.4 (0.8)	287.0 (2.3)	285.9 (2.1)	284.1 (1.2)
75	310.9 (1.2)	310.8 (1.0)	310.8 (1.0)	312.3 (1.0)	312.0 (3.4)	313.2 (2.1)	311.9 (1.2)
90	336.1 (2.0)	335.9 (1.4)	335.9 (1.4)	335.3 (1.2)	333.4 (2.1)	338.4 (2.3)	335.2 (1.3)
95	350.8 (1.7)	350.3 (1.9)	350.3 (1.9)	348.8 (1.6)	345.6 (4.2)	351.9 (1.6)	348.3 (1.2)
FEMALE STUDENTS							
MEAN	291.3 (1.3)	291.2 (1.0)	291.2 (1.0)	294.0 (0.9)	293.8 (1.5)	296.5 (1.2)	289.2 (1.2)
ST. DEV.	44.5 (0.6)	42.2 (0.8)	42.2 (0.8)	39.0 (0.4)	36.3 (0.9)	38.8 (0.8)	40.5 (0.7)
PERCENTILES							
5	215.0 (1.9)	218.9 (2.7)	218.9 (2.7)	227.4 (1.9)	231.7 (3.3)	232.3 (3.8)	219.4 (2.1)
10	233.3 (1.6)	236.8 (2.0)	236.8 (2.0)	242.9 (1.2)	246.5 (4.8)	247.0 (2.1)	236.8 (1.6)
25	262.7 (1.7)	264.9 (1.4)	264.9 (1.4)	268.6 (1.3)	270.2 (2.1)	270.5 (2.3)	262.9 (1.8)
50	293.6 (1.2)	293.4 (0.9)	293.4 (0.9)	295.2 (1.0)	294.6 (2.2)	296.6 (1.2)	290.7 (1.1)
75	321.7 (1.6)	319.7 (0.7)	319.7 (0.7)	320.9 (0.9)	319.4 (1.5)	323.5 (1.5)	317.0 (1.6)
90	346.2 (1.6)	343.3 (1.0)	343.3 (1.0)	343.1 (1.0)	339.8 (1.7)	346.3 (2.5)	339.7 (1.7)
95	360.7 (1.2)	357.0 (1.3)	357.0 (1.3)	355.5 (1.2)	351.7 (2.8)	359.4 (2.7)	353.2 (1.8)

Table 16-84 (continued)
Percentiles for Reading Long-term Trend, Grade 12

	1971	1975	1980	1984	1988	1990	1992
WHITE STUDENTS							
MEAN	291.4 (1.0)	293.0 (0.6)	293.0 (0.6)	295.3 (0.9)	294.7 (1.2)	296.6 (1.2)	292.8 (0.9)
ST. DEV.	42.5 (0.4)	39.8 (0.4)	39.8 (0.4)	38.2 (0.3)	36.0 (0.8)	39.6 (0.6)	37.9 (0.4)
PERCENTILES							
5	219.4 (1.4)	225.9 (1.2)	225.9 (1.2)	229.9 (1.4)	232.6 (1.1)	228.5 (2.5)	228.5 (1.4)
10	236.6 (1.0)	241.7 (0.9)	241.7 (0.9)	245.6 (0.9)	247.3 (3.7)	246.2 (2.5)	243.5 (1.5)
25	263.9 (1.4)	267.0 (0.9)	267.0 (0.9)	270.7 (1.1)	271.4 (1.7)	271.1 (1.4)	267.7 (1.0)
50	292.9 (1.2)	294.0 (0.8)	294.0 (0.8)	296.7 (1.1)	295.4 (1.6)	297.5 (1.2)	293.6 (0.8)
75	320.1 (1.1)	319.9 (0.7)	319.9 (0.7)	321.6 (0.8)	319.9 (1.9)	323.8 (1.9)	318.8 (1.0)
90	344.5 (1.0)	343.2 (0.7)	343.2 (0.7)	343.2 (0.8)	339.7 (1.6)	347.1 (1.6)	340.6 (1.3)
95	358.9 (1.4)	357.0 (1.2)	357.0 (1.2)	355.8 (0.9)	351.6 (3.0)	359.7 (1.7)	353.5 (1.4)
BLACK STUDENTS							
MEAN	238.7 (1.7)	240.6 (2.0)	240.6 (2.0)	263.6 (1.2)	274.4 (2.4)	267.3 (2.3)	243.1 (1.8)
ST. DEV.	43.5 (0.7)	43.8 (1.2)	43.8 (1.2)	37.0 (0.8)	35.9 (1.3)	39.2 (2.2)	39.5 (1.2)
PERCENTILES							
5	164.7 (4.4)	164.7 (3.1)	164.7 (3.1)	201.9 (4.1)	214.4 (9.6)	201.3 (7.9)	176.0 (2.4)
10	182.1 (4.2)	182.4 (5.3)	182.4 (5.3)	216.0 (2.0)	227.8 (4.3)	217.4 (4.0)	191.1 (3.6)
25	210.4 (2.4)	212.1 (3.0)	212.1 (3.0)	239.0 (1.4)	250.5 (2.5)	242.4 (3.9)	217.0 (2.7)
50	239.3 (1.6)	242.1 (1.6)	242.1 (1.6)	264.2 (1.2)	274.3 (3.6)	268.4 (1.9)	243.9 (2.6)
75	268.1 (2.0)	271.6 (1.4)	271.6 (1.4)	288.3 (1.6)	299.6 (3.1)	293.7 (2.7)	270.1 (2.0)
90	294.1 (2.4)	295.7 (1.4)	295.7 (1.4)	310.5 (1.9)	321.0 (4.0)	316.2 (4.8)	293.3 (1.7)
95	309.7 (2.2)	308.3 (2.7)	308.3 (2.7)	323.6 (3.4)	333.1 (4.9)	330.5 (11.0)	306.6 (2.4)
HISPANIC STUDENTS *							
MEAN	0.0 (0.0)	252.4 (3.6)	252.4 (3.6)	268.1 (2.9)	270.8 (4.3)	274.8 (3.6)	261.4 (2.7)
ST. DEV.	0.0 (0.0)	42.0 (2.2)	42.0 (2.2)	39.7 (1.5)	37.7 (2.0)	40.7 (2.7)	40.1 (1.4)
PERCENTILES							
5	0.0 (0.0)	184.4 (3.7)	184.4 (3.7)	201.5 (2.4)	204.2 (11.7)	205.9 (11.1)	194.3 (7.8)
10	0.0 (0.0)	197.1 (4.9)	197.1 (4.9)	216.6 (2.9)	218.0 (7.4)	224.3 (12.0)	208.2 (3.7)
25	0.0 (0.0)	225.4 (5.9)	225.4 (5.9)	241.5 (2.6)	246.4 (5.9)	250.4 (8.3)	235.3 (5.0)
50	0.0 (0.0)	252.8 (3.7)	252.8 (3.7)	268.6 (3.1)	273.6 (5.1)	276.3 (3.2)	262.6 (3.5)
75	0.0 (0.0)	279.4 (3.0)	279.4 (3.0)	295.4 (3.9)	297.9 (7.1)	302.6 (4.9)	288.6 (3.2)
90	0.0 (0.0)	306.7 (6.1)	306.7 (6.1)	318.3 (6.1)	315.9 (18.1)	326.5 (3.2)	312.6 (3.0)
95	0.0 (0.0)	320.8 (6.8)	320.8 (6.8)	332.3 (7.7)	328.0 (8.6)	339.4 (11.2)	325.1 (3.4)

* NO DATA WERE AVAILABLE FOR HISPANIC STUDENTS IN 1971

Table 16-85

Weighted Means, Standard Deviations, and Percentiles of Mathematics Distributions for Mathematics Main Samples, by Subgroup, Grade 4

	MEAN	STD DEV	5TH	10TH	25TH	50TH	75TH	90TH	95TH
-- TOTAL --	218.5(0.7)	32.4(0.4)	162.4(1.1)	175.3(0.9)	197.1(1.0)	220.1(1.0)	241.1(1.2)	258.9(0.8)	269.4(1.5)
GENDER									
MALE	219.7(0.8)	33.3(0.5)	162.0(1.9)	175.0(1.4)	197.8(1.2)	221.6(1.0)	242.9(1.5)	260.9(1.6)	271.8(1.6)
FEMALE	217.3(1.0)	31.5(0.5)	162.8(1.8)	175.5(1.1)	196.5(1.0)	218.7(1.5)	239.3(1.3)	256.7(1.1)	266.5(0.7)
RACE/ETHNICITY									
WHITE	226.8(0.9)	29.1(0.4)	176.7(1.8)	188.6(1.4)	207.8(1.3)	228.1(0.8)	246.5(1.1)	262.7(1.1)	272.5(1.2)
BLACK	191.5(1.3)	28.8(0.8)	143.6(2.4)	154.0(3.4)	171.8(1.8)	191.8(1.7)	210.4(1.8)	227.9(2.7)	238.4(1.5)
HISPANIC	200.8(1.4)	30.2(0.7)	149.6(2.2)	161.7(2.1)	180.7(1.5)	201.5(1.6)	221.5(2.6)	239.7(2.4)	249.8(3.3)
ASIAN/PACIFIC AMERICAN	231.3(2.4)	31.4(1.5)	173.2(7.6)	188.1(4.8)	212.1(8.3)	233.2(3.8)	253.6(5.7)	270.5(4.3)	280.5(10.2)
AMER IND/ALASKAN NATV	209.1(3.2)	31.0(2.0)	158.3(10.2)	169.9(5.8)	191.2(3.3)	208.4(2.6)	227.7(5.4)	247.5(6.0)	264.1(21.6)
PARENTS' EDUCATION LEVEL									
LESS THAN H.S.	203.9(2.6)	29.5(1.9)	153.9(5.8)	164.8(9.2)	185.1(3.9)	205.0(2.8)	224.7(1.9)	241.6(3.9)	250.0(6.2)
GRADUATED H.S.	213.1(1.5)	30.5(0.9)	161.1(3.2)	172.8(2.7)	192.6(2.8)	215.4(1.3)	234.2(1.6)	251.3(2.7)	262.1(2.8)
SOME EDUC AFTER H.S.	223.5(1.5)	31.2(1.3)	166.0(11.5)	181.7(3.0)	204.3(3.0)	227.2(1.5)	245.1(2.1)	259.4(4.9)	268.8(5.2)
GRADUATED COLLEGE	225.6(1.0)	33.1(0.6)	166.4(2.0)	181.3(2.1)	204.5(1.2)	227.9(1.2)	248.4(1.5)	266.2(1.6)	276.4(2.1)
UNKNOWN	212.8(0.8)	30.9(0.6)	160.5(2.0)	172.0(1.3)	192.3(0.8)	213.5(0.9)	234.6(1.4)	252.2(1.4)	261.3(1.3)

Table 16-86

Weighted Means, Standard Deviations, and Percentiles of Mathematics Distributions
for Mathematics Main Samples, by Subgroup, Grade 8

	MEAN	STD DEV	5TH	10TH	25TH	50TH	75TH	90TH	95TH
-- TOTAL --	267.7(0.9)	36.6(0.5)	206.2(1.9)	219.5(0.9)	242.3(1.3)	268.5(1.4)	293.8(0.9)	314.6(1.0)	326.3(1.8)
GENDER									
MALE	267.4(1.1)	37.2(0.6)	205.3(2.8)	218.2(2.1)	241.6(1.4)	268.1(1.3)	294.3(1.2)	314.7(1.5)	326.8(2.1)
FEMALE	268.0(1.0)	36.0(0.5)	207.5(2.9)	220.9(1.3)	242.9(2.5)	268.9(1.3)	293.3(1.5)	314.5(1.0)	325.9(2.2)
RACE/ETHNICITY									
WHITE	277.2(1.0)	32.8(0.4)	222.1(1.6)	234.0(1.3)	254.7(1.4)	278.2(1.3)	300.1(1.1)	319.0(1.1)	329.6(1.4)
BLACK	236.8(1.4)	30.8(1.0)	186.9(2.9)	197.7(2.5)	215.8(1.4)	236.6(1.5)	257.8(1.5)	275.9(2.2)	287.1(3.5)
HISPANIC	246.3(1.2)	34.8(0.8)	189.6(4.2)	202.3(2.8)	222.1(1.2)	246.4(1.7)	270.0(1.8)	291.0(2.0)	303.2(2.1)
ASIAN/PACIFIC AMERICAN	288.0(5.5)	37.3(1.6)	226.1(3.2)	240.3(8.6)	262.3(5.1)	287.8(11.6)	314.3(5.4)	339.3(9.5)	347.0(3.4)
AMER IND/ALASKAN NATV	254.3(2.8)	27.8(1.4)	207.4(7.2)	219.9(3.6)	235.0(2.5)	253.7(4.0)	274.1(6.2)	291.8(3.7)	302.6(6.9)
PARENTS' EDUCATION LEVEL									
LESS THAN H.S.	248.5(1.7)	32.0(1.2)	198.8(2.4)	207.9(2.1)	226.3(1.4)	246.0(3.8)	270.8(2.7)	290.7(3.1)	301.9(5.2)
GRADUATED H.S.	256.6(1.2)	33.3(0.8)	201.2(3.5)	212.8(1.7)	233.7(1.9)	257.3(1.6)	280.5(1.9)	298.5(1.3)	309.9(1.9)
SOME EDUC AFTER H.S.	270.4(1.1)	33.6(1.0)	215.2(3.5)	227.0(2.6)	248.6(1.6)	270.3(2.1)	293.3(0.9)	313.9(1.6)	325.3(2.2)
GRADUATED COLLEGE	280.0(1.2)	36.1(0.7)	216.2(1.7)	231.2(2.4)	256.4(1.0)	283.0(1.6)	305.8(1.5)	324.0(1.1)	334.3(1.5)
UNKNOWN	251.0(1.6)	34.1(1.0)	193.4(2.9)	206.4(3.8)	227.4(4.0)	250.2(2.8)	274.1(4.1)	296.1(2.2)	307.7(2.2)

BEST COPY AVAILABLE

Table 16-87

Weighted Means, Standard Deviations, and Percentiles of Mathematics Distributions for Mathematics Main Samples, by Subgroup, Grade 12

	MEAN	STD DEV	5TH	10TH	25TH	50TH	75TH	90TH	95TH
-- TOTAL --	298.7(0.9)	34.4(0.4)	240.3(1.9)	252.9(1.2)	275.0(1.4)	299.9(1.2)	323.2(1.3)	342.6(1.0)	353.6(1.3)
GENDER									
MALE	300.6(1.1)	35.2(0.6)	240.7(2.5)	253.2(1.5)	276.5(2.0)	301.5(1.7)	326.0(1.5)	345.5(1.1)	356.7(2.0)
FEMALE	297.0(1.0)	33.5(0.5)	239.9(2.9)	252.4(1.3)	273.7(1.0)	298.3(1.0)	320.4(1.5)	339.7(1.3)	350.6(1.1)
RACE/ETHNICITY									
WHITE	305.0(0.9)	32.3(0.5)	249.9(1.7)	262.0(0.7)	283.4(1.5)	306.7(1.3)	328.0(1.4)	345.6(1.1)	356.0(1.0)
BLACK	274.8(1.7)	30.4(0.8)	226.6(1.8)	235.0(2.7)	252.6(2.0)	273.7(2.4)	296.5(2.6)	314.9(1.8)	326.1(2.2)
HISPANIC	282.9(1.8)	32.9(1.3)	229.1(18.1)	240.5(4.0)	260.6(3.3)	282.8(1.8)	304.5(1.7)	324.8(3.0)	339.6(1.6)
ASIAN/PACIFIC AMERICAN	315.3(3.5)	32.0(1.6)	259.2(5.6)	274.3(6.0)	293.8(3.6)	315.2(3.8)	340.1(7.6)	357.6(3.5)	366.0(5.7)
AMER IND/ALASKAN NATV	281.1(9.0)	34.1(4.6)	223.2(27.8)	232.5(19.1)	255.5(29.8)	281.0(14.3)	311.3(11.6)	322.3(2.7)	333.1(12.9)
PARENTS' EDUCATION LEVEL									
LESS THAN H.S.	277.5(1.7)	30.4(1.1)	226.8(3.2)	238.2(2.1)	258.2(2.7)	278.0(2.1)	297.7(2.3)	316.8(2.5)	329.1(3.6)
GRADUATED H.S.	287.0(1.4)	31.5(0.7)	234.6(2.6)	245.8(1.6)	265.2(2.3)	287.6(1.6)	310.0(2.8)	328.0(1.4)	337.7(2.1)
SOME EDUC AFTER H.S.	297.8(1.0)	32.1(0.6)	242.7(2.9)	255.5(1.2)	275.7(1.7)	298.8(0.8)	320.2(1.7)	339.3(1.1)	349.7(1.5)
GRADUATED COLLEGE	310.1(1.2)	33.2(0.7)	251.2(3.2)	264.1(1.6)	288.9(2.2)	312.2(1.8)	333.9(1.2)	351.6(1.4)	361.3(1.8)
UNKNOWN	275.6(3.0)	32.0(3.1)	225.8(6.3)	235.6(3.6)	251.3(6.6)	274.7(4.1)	299.6(13.4)	317.4(3.4)	327.1(2.3)

Table 16-88

Percentiles for Mathematics Long-term Trend, Grade 4

	1978	1982	1986	1990	1992
TOTAL SAMPLE					
MEAN	218.6 (0.8)	219.0 (1.1)	221.7 (1.0)	229.6 (0.8)	229.6 (0.8)
ST. DEV.	36.0 (0.3)	34.8 (0.4)	34.0 (0.5)	32.9 (0.5)	33.1 (0.5)
PERCENTILES					
5	157.1 (1.0)	159.3 (1.8)	163.0 (1.3)	173.3 (2.6)	172.2 (1.6)
10	171.1 (1.2)	173.2 (1.8)	176.7 (1.5)	185.8 (2.2)	185.4 (1.2)
25	194.6 (1.0)	196.0 (1.1)	199.0 (1.6)	207.8 (1.3)	207.9 (1.2)
50	220.1 (1.0)	220.4 (1.2)	223.3 (1.1)	231.1 (0.9)	231.0 (0.8)
75	243.7 (0.9)	243.3 (1.4)	245.6 (1.2)	252.5 (0.7)	252.6 (0.8)
90	264.0 (1.2)	262.7 (1.0)	264.2 (1.3)	271.0 (1.0)	270.9 (1.3)
95	275.7 (1.2)	273.8 (1.3)	275.5 (1.2)	282.1 (1.3)	281.7 (1.2)
MALE STUDENTS					
MEAN	217.4 (0.7)	217.1 (1.2)	221.7 (1.1)	229.1 (0.9)	230.8 (1.0)
ST. DEV.	36.7 (0.5)	35.8 (0.5)	34.3 (0.8)	33.5 (0.6)	33.5 (0.6)
PERCENTILES					
5	154.9 (2.3)	156.4 (2.1)	162.7 (2.0)	171.8 (2.5)	172.7 (1.5)
10	169.0 (1.3)	170.2 (1.4)	176.1 (1.7)	184.6 (2.1)	186.1 (1.4)
25	192.8 (1.0)	193.0 (1.5)	198.6 (1.6)	206.7 (1.2)	208.9 (1.6)
50	218.4 (0.9)	218.6 (1.7)	223.0 (1.0)	230.4 (1.0)	232.2 (1.0)
75	243.0 (1.1)	242.3 (1.6)	245.7 (1.6)	252.4 (0.8)	254.2 (1.1)
90	263.8 (1.2)	262.2 (1.2)	265.1 (1.9)	271.6 (1.8)	272.5 (1.2)
95	275.2 (1.1)	273.6 (1.9)	276.4 (2.1)	282.8 (1.7)	283.8 (1.4)
FEMALE STUDENTS					
MEAN	219.9 (1.0)	220.8 (1.2)	221.7 (1.2)	230.2 (1.1)	228.4 (1.0)
ST. DEV.	35.3 (0.4)	33.7 (0.5)	33.7 (0.6)	32.4 (0.6)	32.7 (0.7)
PERCENTILES					
5	159.4 (1.3)	162.8 (1.7)	163.5 (2.3)	174.5 (2.8)	171.8 (1.6)
10	173.1 (2.0)	176.6 (1.6)	177.5 (2.6)	187.0 (2.7)	184.9 (1.6)
25	196.4 (1.2)	198.9 (1.8)	199.2 (1.8)	208.9 (1.3)	206.9 (1.5)
50	221.5 (1.0)	222.2 (1.1)	223.5 (1.1)	231.8 (1.0)	229.9 (1.0)
75	244.3 (1.5)	244.2 (1.4)	245.5 (1.5)	252.7 (1.0)	251.1 (1.1)
90	264.2 (1.4)	263.1 (1.0)	263.3 (1.6)	270.4 (1.3)	269.2 (1.2)
95	276.1 (1.8)	273.9 (1.7)	274.2 (2.0)	281.4 (1.1)	279.8 (2.3)

Table 16-88 (continued)
Percentiles for Mathematics Long-term Trend, Grade 4

	1978	1982	1986	1990	1992
WHITE STUDENTS					
MEAN	224.1 (0.9)	224.0 (1.1)	226.9 (1.1)	235.2 (0.8)	235.1 (0.8)
ST. DEV.	34.0 (0.3)	32.8 (0.4)	32.6 (0.5)	31.2 (0.5)	31.0 (0.5)
PERCENTILES					
5	166.3 (1.5)	168.1 (1.4)	170.6 (2.4)	181.8 (2.4)	181.8 (1.5)
10	179.4 (1.5)	180.8 (1.7)	183.9 (1.7)	194.0 (1.6)	194.2 (1.5)
25	201.4 (1.1)	201.9 (1.3)	205.3 (1.1)	214.6 (0.9)	215.0 (0.9)
50	225.1 (1.0)	225.3 (1.4)	228.3 (1.1)	236.3 (1.0)	236.1 (1.1)
75	247.7 (0.8)	246.8 (0.9)	249.6 (0.8)	256.4 (0.6)	256.4 (1.0)
90	267.0 (1.1)	265.3 (1.0)	267.4 (1.2)	274.5 (0.8)	273.9 (1.3)
95	278.4 (1.7)	276.0 (1.3)	278.2 (1.8)	284.8 (2.1)	284.5 (1.6)
BLACK STUDENTS					
MEAN	192.4 (1.1)	194.9 (1.6)	201.6 (1.6)	208.4 (2.2)	208.0 (2.0)
ST. DEV.	34.5 (0.8)	33.7 (0.8)	31.7 (1.1)	31.5 (0.8)	31.8 (0.7)
PERCENTILES					
5	133.7 (1.9)	136.7 (2.5)	146.2 (3.2)	156.0 (1.7)	154.9 (3.4)
10	147.0 (1.7)	150.4 (2.3)	158.4 (4.9)	167.1 (3.7)	165.9 (2.9)
25	169.3 (1.9)	172.5 (2.0)	180.5 (4.1)	186.0 (4.1)	185.5 (2.4)
50	193.0 (1.1)	196.6 (2.0)	202.9 (1.6)	208.4 (3.1)	208.6 (2.1)
75	216.4 (1.6)	218.2 (2.0)	223.6 (2.0)	231.4 (2.1)	230.4 (2.0)
90	236.1 (1.6)	236.7 (2.5)	241.2 (1.7)	248.9 (2.9)	249.2 (2.1)
95	247.5 (1.4)	247.9 (2.8)	251.3 (1.3)	258.9 (4.3)	258.7 (3.4)
HISPANIC STUDENTS					
MEAN	202.9 (2.2)	204.0 (1.3)	205.4 (2.1)	213.8 (2.1)	211.9 (2.3)
ST. DEV.	35.1 (1.4)	32.8 (1.1)	31.1 (1.9)	30.3 (1.2)	31.9 (1.4)
PERCENTILES					
5	144.4 (5.4)	148.1 (2.8)	154.8 (3.7)	161.8 (3.4)	158.6 (4.4)
10	156.3 (3.7)	160.8 (3.2)	163.8 (1.8)	173.4 (1.4)	169.0 (3.5)
25	178.7 (3.2)	181.3 (2.3)	184.6 (3.2)	193.1 (3.6)	189.7 (2.2)
50	204.3 (3.0)	205.2 (1.6)	206.3 (2.4)	216.2 (4.1)	211.8 (3.5)
75	227.2 (2.5)	226.5 (2.0)	226.0 (3.8)	235.1 (3.3)	233.8 (3.4)
90	249.5 (4.0)	246.4 (3.4)	244.8 (3.8)	251.7 (3.4)	252.7 (3.8)
95	259.6 (4.6)	256.6 (2.9)	254.4 (4.6)	262.2 (3.5)	263.1 (6.8)

Table 16-89

Percentiles for Mathematics Long-term Trend, Grade 8

	1978	1982	1986	1990	1992
TOTAL SAMPLE					
MEAN	264.1 (1.1)	268.6 (1.1)	269.0 (1.2)	270.4 (0.9)	273.1 (0.9)
ST. DEV.	39.0 (0.5)	33.4 (0.5)	30.8 (0.5)	31.1 (0.5)	30.9 (0.6)
PERCENTILES					
5	198.2 (1.6)	212.4 (2.7)	218.3 (1.8)	217.6 (2.2)	220.5 (2.0)
10	213.3 (1.5)	225.3 (1.6)	230.0 (1.4)	230.2 (1.4)	233.2 (1.2)
25	238.1 (1.3)	246.2 (1.2)	248.3 (1.8)	249.8 (0.9)	252.9 (1.1)
50	265.2 (1.1)	269.5 (1.0)	268.7 (1.3)	270.9 (1.0)	274.1 (0.7)
75	291.1 (1.1)	291.6 (1.1)	289.6 (1.3)	291.7 (1.0)	294.0 (1.0)
90	313.4 (1.2)	310.8 (1.2)	309.2 (1.5)	309.9 (1.0)	311.9 (1.6)
95	326.6 (1.3)	322.2 (1.2)	320.5 (2.2)	320.1 (1.6)	322.9 (1.2)
MALE STUDENTS					
MEAN	263.6 (1.3)	269.2 (1.4)	270.0 (1.1)	271.2 (1.2)	274.1 (1.1)
ST. DEV.	40.1 (0.5)	34.4 (0.7)	31.6 (0.7)	32.4 (0.7)	31.6 (0.9)
PERCENTILES					
5	195.8 (1.4)	211.5 (2.2)	218.0 (1.8)	215.5 (2.1)	220.5 (2.9)
10	211.4 (1.4)	224.3 (2.0)	229.5 (1.7)	228.6 (2.0)	233.2 (2.0)
25	236.7 (1.4)	246.1 (1.5)	248.9 (2.3)	250.2 (1.7)	253.1 (1.8)
50	264.8 (1.4)	270.2 (1.2)	270.1 (1.6)	272.0 (1.0)	274.9 (1.0)
75	291.5 (1.5)	293.3 (1.2)	291.4 (1.6)	293.1 (1.2)	295.7 (0.8)
90	314.4 (1.7)	312.5 (1.5)	310.8 (1.5)	312.4 (1.4)	314.0 (1.6)
95	327.5 (1.5)	324.1 (1.3)	322.0 (2.6)	323.1 (1.9)	324.8 (2.1)
FEMALE STUDENTS					
MEAN	264.7 (1.1)	268.0 (1.1)	267.9 (1.5)	269.6 (0.9)	272.0 (1.0)
ST. DEV.	37.9 (0.6)	32.3 (0.5)	30.0 (0.7)	29.7 (0.5)	30.3 (0.6)
PERCENTILES					
5	200.9 (2.6)	213.5 (1.5)	218.5 (3.2)	220.4 (2.3)	220.6 (1.0)
10	215.0 (1.6)	226.2 (1.4)	230.6 (2.0)	231.4 (1.2)	233.0 (1.3)
25	239.4 (1.4)	246.3 (1.1)	247.8 (1.6)	249.5 (1.1)	252.7 (1.2)
50	265.7 (1.2)	268.8 (0.9)	267.4 (1.7)	269.9 (1.2)	273.4 (1.0)
75	290.7 (1.0)	290.1 (1.1)	287.8 (1.7)	290.3 (1.3)	292.2 (1.3)
90	312.4 (1.4)	308.8 (1.5)	307.2 (2.8)	307.7 (1.5)	309.8 (1.2)
95	325.6 (1.2)	320.1 (2.0)	318.5 (2.4)	317.3 (0.8)	320.8 (1.1)

Table 16-89 (continued)
Percentiles for Mathematics Long-term Trend, Grade 8

	1978	1982	1986	1990	1992
WHITE STUDENTS					
MEAN	271.6 (0.8)	274.4 (1.0)	273.6 (1.3)	276.3 (1.1)	278.9 (0.9)
ST. DEV.	35.7 (0.5)	31.0 (0.4)	29.4 (0.6)	29.0 (0.5)	28.5 (0.5)
PERCENTILES					
5	211.9 (1.4)	223.0 (1.6)	225.7 (1.5)	228.2 (1.5)	230.9 (1.6)
10	225.5 (1.4)	234.4 (1.2)	236.5 (1.3)	239.3 (1.0)	242.2 (1.4)
25	247.6 (0.9)	253.5 (1.1)	254.1 (1.4)	257.3 (1.1)	260.5 (0.8)
50	272.2 (1.0)	274.9 (0.9)	273.3 (1.0)	276.6 (1.0)	279.4 (1.0)
75	296.0 (0.7)	295.5 (1.0)	293.2 (1.3)	296.0 (1.1)	298.0 (1.1)
90	317.1 (1.2)	313.8 (1.4)	312.1 (2.2)	313.2 (1.3)	315.1 (1.3)
95	329.6 (1.3)	324.8 (1.4)	322.9 (1.8)	322.9 (1.6)	325.2 (1.4)
BLACK STUDENTS					
MEAN	229.6 (1.9)	240.4 (1.6)	249.2 (2.3)	249.1 (2.3)	250.2 (1.9)
ST. DEV.	36.0 (0.6)	31.0 (1.1)	28.3 (1.1)	28.7 (1.2)	30.1 (1.2)
PERCENTILES					
5	170.2 (1.9)	189.0 (4.3)	201.7 (4.5)	201.6 (5.4)	199.5 (4.5)
10	184.1 (2.6)	200.2 (3.7)	213.2 (2.3)	211.8 (2.2)	212.3 (5.1)
25	205.5 (1.9)	219.3 (1.8)	230.7 (2.2)	229.9 (3.0)	231.1 (3.0)
50	229.0 (2.2)	241.0 (1.9)	249.3 (2.3)	249.4 (2.0)	250.6 (1.9)
75	254.1 (2.2)	260.9 (1.4)	266.9 (1.5)	267.8 (2.9)	270.9 (1.8)
90	276.4 (2.4)	279.7 (2.2)	284.4 (3.7)	285.3 (2.8)	286.5 (2.1)
95	288.4 (3.9)	291.1 (1.7)	296.4 (4.3)	296.2 (4.1)	297.4 (3.5)
HISPANIC STUDENTS					
MEAN	238.0 (2.0)	252.4 (1.7)	254.3 (2.9)	254.6 (1.8)	259.3 (1.8)
ST. DEV.	35.2 (1.1)	31.0 (1.0)	29.3 (1.3)	29.9 (1.2)	28.1 (1.0)
PERCENTILES					
5	180.2 (1.8)	202.3 (2.2)	205.9 (3.6)	206.2 (3.7)	212.2 (3.5)
10	192.5 (2.2)	213.5 (2.6)	216.2 (3.8)	216.4 (3.1)	224.0 (2.4)
25	214.3 (1.8)	230.7 (1.9)	235.5 (2.7)	234.3 (2.2)	240.6 (3.2)
50	237.4 (2.0)	251.9 (1.4)	254.3 (3.4)	255.1 (1.9)	259.4 (2.3)
75	261.9 (3.2)	273.7 (1.4)	274.2 (2.4)	275.2 (3.5)	278.6 (2.9)
90	283.7 (3.4)	292.8 (2.4)	291.7 (3.1)	292.2 (2.9)	294.9 (1.6)
95	296.3 (3.1)	304.1 (2.9)	301.2 (1.9)	303.3 (3.3)	304.1 (3.2)

Table 16-90

Percentiles for Mathematics Long-term Trend, Grade 12

	1978	1982	1986	1990	1992
TOTAL SAMPLE					
MEAN	300.4 (1.0)	298.5 (0.9)	302.0 (0.9)	304.6 (0.9)	306.7 (0.9)
ST. DEV.	34.9 (0.3)	32.4 (0.4)	31.0 (0.5)	31.1 (0.6)	30.1 (0.5)
PERCENTILES					
5	241.3 (1.3)	244.9 (1.1)	251.7 (1.2)	253.4 (1.0)	255.6 (2.1)
10	254.2 (1.1)	255.9 (1.0)	262.7 (1.0)	264.0 (1.1)	267.2 (1.6)
25	276.4 (1.2)	275.8 (1.3)	280.7 (0.6)	282.5 (1.0)	286.3 (1.2)
50	301.4 (1.1)	298.8 (1.0)	301.4 (1.3)	304.9 (1.1)	307.6 (1.0)
75	325.4 (1.0)	321.5 (0.8)	323.1 (1.9)	326.5 (1.2)	328.0 (1.0)
90	344.7 (0.8)	340.6 (0.9)	343.0 (1.3)	344.5 (1.3)	345.2 (1.1)
95	355.7 (0.9)	351.2 (1.1)	354.0 (1.1)	355.5 (2.2)	354.8 (1.0)
MALE STUDENTS					
MEAN	303.8 (1.0)	301.5 (1.0)	304.7 (1.2)	306.3 (1.1)	308.9 (1.1)
ST. DEV.	35.4 (0.4)	32.8 (0.5)	32.0 (0.7)	32.3 (0.7)	30.8 (0.6)
PERCENTILES					
5	243.8 (1.2)	247.0 (1.3)	252.7 (3.0)	252.8 (3.0)	257.8 (1.7)
10	257.0 (1.2)	257.9 (1.2)	264.1 (1.2)	263.9 (1.2)	268.9 (1.8)
25	278.9 (1.2)	278.1 (1.1)	282.3 (1.8)	283.7 (1.3)	287.8 (1.2)
50	304.8 (1.3)	301.8 (1.6)	303.9 (1.2)	306.4 (1.6)	309.0 (1.6)
75	329.5 (1.1)	325.1 (1.2)	327.8 (2.1)	329.3 (1.1)	331.4 (1.1)
90	349.2 (1.0)	344.4 (1.1)	346.7 (1.6)	347.8 (1.4)	348.6 (1.3)
95	360.1 (1.0)	354.4 (1.8)	357.5 (1.7)	358.5 (1.3)	358.1 (1.4)
FEMALE STUDENTS					
MEAN	297.1 (1.0)	295.6 (1.0)	299.4 (1.0)	302.9 (1.1)	304.5 (1.1)
ST. DEV.	34.0 (0.4)	31.7 (0.4)	29.9 (0.7)	29.9 (0.9)	29.3 (0.7)
PERCENTILES					
5	239.3 (1.3)	242.8 (1.6)	250.3 (2.8)	253.9 (1.9)	253.7 (2.3)
10	252.2 (1.0)	254.1 (1.2)	261.2 (1.4)	264.0 (1.5)	265.6 (2.4)
25	274.3 (1.3)	273.7 (1.2)	279.3 (1.3)	281.5 (1.3)	284.8 (1.1)
50	298.3 (1.1)	296.1 (1.2)	299.1 (1.3)	303.7 (1.7)	305.8 (1.5)
75	321.5 (1.0)	317.7 (0.8)	319.8 (1.7)	324.1 (1.2)	324.8 (1.2)
90	340.3 (1.4)	336.7 (1.7)	338.2 (2.2)	341.4 (1.6)	341.4 (2.1)
95	350.4 (1.5)	347.2 (1.5)	349.3 (1.9)	351.8 (2.2)	350.6 (2.3)

Table 16-90 (continued)

Percentiles for Mathematics Long-term Trend, Grade 12

	1978	1982	1986	1990	1992
WHITE STUDENTS					
MEAN	305.9 (0.9)	303.7 (0.9)	307.5 (1.0)	309.5 (1.0)	311.9 (0.8)
ST. DEV.	32.3 (0.2)	30.4 (0.4)	29.1 (0.6)	29.5 (0.5)	28.4 (0.5)
PERCENTILES					
5	251.9 (0.6)	253.3 (1.1)	261.2 (1.6)	260.2 (1.3)	264.1 (2.0)
10	263.3 (1.3)	263.8 (1.1)	270.5 (1.3)	270.5 (1.5)	274.4 (1.4)
25	283.5 (1.0)	282.3 (1.1)	286.9 (1.2)	288.8 (1.5)	292.8 (1.1)
50	306.6 (1.0)	303.9 (1.2)	306.8 (1.3)	310.1 (1.3)	312.8 (1.0)
75	328.9 (0.8)	325.1 (0.9)	327.8 (1.7)	330.1 (1.2)	332.2 (1.0)
90	347.3 (0.7)	343.4 (1.1)	346.1 (1.3)	347.2 (1.0)	348.0 (1.0)
95	357.8 (0.7)	353.4 (1.5)	356.0 (1.4)	357.1 (1.3)	357.4 (1.2)
BLACK STUDENTS					
MEAN	268.4 (1.3)	271.8 (1.2)	278.6 (2.1)	288.5 (2.8)	285.8 (2.2)
ST. DEV.	31.8 (1.0)	29.2 (0.7)	26.4 (1.4)	27.9 (1.7)	27.5 (1.3)
PERCENTILES					
5	217.2 (2.0)	225.1 (1.4)	236.7 (3.9)	245.4 (4.4)	238.5 (4.3)
10	227.8 (1.7)	234.5 (1.7)	244.3 (4.2)	253.5 (3.5)	248.9 (6.9)
25	245.7 (1.2)	251.4 (1.6)	259.9 (1.6)	268.7 (1.8)	267.4 (3.8)
50	267.7 (1.6)	271.2 (1.4)	278.6 (3.9)	287.1 (2.5)	286.9 (1.9)
75	290.5 (2.2)	291.2 (1.7)	296.1 (2.5)	307.1 (5.3)	303.9 (3.9)
90	310.3 (2.1)	310.8 (1.7)	312.0 (7.4)	325.7 (5.8)	320.8 (2.3)
95	320.7 (2.5)	321.3 (2.2)	324.8 (4.1)	337.7 (4.2)	330.8 (3.0)
HISPANIC STUDENTS					
MEAN	276.3 (2.3)	276.7 (1.8)	283.1 (2.9)	283.5 (2.9)	292.2 (2.6)
ST. DEV.	32.9 (1.0)	29.3 (1.0)	28.7 (2.0)	31.8 (1.8)	26.9 (1.0)
PERCENTILES					
5	224.1 (4.4)	232.0 (1.7)	236.3 (5.3)	229.1 (5.4)	247.5 (4.3)
10	234.0 (2.9)	240.7 (3.2)	248.5 (4.5)	242.2 (8.1)	257.8 (3.5)
25	253.4 (1.8)	255.8 (2.4)	264.7 (2.8)	263.8 (6.8)	273.3 (4.5)
50	275.1 (3.6)	275.3 (3.2)	283.1 (2.5)	281.8 (2.4)	291.6 (3.4)
75	298.5 (3.9)	297.1 (2.6)	301.2 (4.2)	304.0 (4.4)	310.7 (3.7)
90	319.5 (3.9)	314.9 (2.6)	318.6 (2.3)	325.1 (3.6)	327.7 (4.8)
95	332.0 (0.9)	326.7 (4.4)	329.3 (7.3)	336.3 (8.6)	336.4 (2.7)

Table 16-91

Percentiles for Science Long-term Trend, Grade 4

	1977	1982	1986	1990	1992
TOTAL SAMPLE					
MEAN	219.9 (1.2)	220.8 (1.8)	224.3 (1.2)	228.7 (0.8)	230.6 (1.0)
ST. DEV.	44.9 (0.6)	40.9 (1.4)	41.6 (0.6)	40.2 (0.4)	39.9 (0.7)
PERCENTILES					
5	143.8 (2.3)	150.9 (4.9)	155.0 (1.3)	159.8 (1.3)	162.8 (2.0)
10	160.9 (2.1)	166.8 (2.6)	169.9 (1.8)	176.1 (1.1)	177.8 (1.8)
25	190.1 (1.6)	194.4 (2.2)	195.9 (1.3)	202.0 (1.4)	203.8 (1.6)
50	221.5 (1.1)	221.4 (2.4)	225.1 (1.7)	230.3 (0.9)	232.1 (0.9)
75	251.0 (1.1)	249.0 (2.0)	253.1 (1.7)	256.6 (0.8)	258.4 (1.0)
90	276.5 (1.2)	272.4 (3.9)	276.9 (2.0)	278.8 (1.3)	280.6 (1.6)
95	291.4 (1.2)	286.4 (3.7)	290.9 (1.9)	292.1 (1.4)	293.6 (1.4)
MALE STUDENTS					
MEAN	222.1 (1.3)	221.0 (2.3)	227.3 (1.4)	230.3 (1.1)	234.7 (1.2)
ST. DEV.	45.0 (0.7)	42.0 (2.0)	41.9 (0.7)	41.9 (0.6)	40.7 (1.0)
PERCENTILES					
5	146.8 (2.6)	150.4 (5.5)	158.0 (3.6)	159.6 (2.2)	164.7 (3.0)
10	163.2 (1.9)	166.5 (3.8)	172.9 (1.8)	176.3 (2.3)	180.9 (2.7)
25	191.9 (1.9)	193.5 (4.1)	198.7 (1.8)	202.1 (2.5)	207.2 (1.9)
50	223.6 (1.4)	221.3 (3.6)	227.9 (1.7)	231.6 (1.9)	236.2 (1.5)
75	253.4 (1.4)	250.4 (3.1)	256.1 (1.9)	259.4 (1.0)	263.1 (1.5)
90	279.1 (1.3)	274.7 (4.3)	280.3 (2.0)	283.3 (1.8)	285.8 (1.5)
95	294.2 (1.5)	287.1 (5.3)	294.8 (2.7)	296.3 (2.4)	298.6 (1.5)
FEMALE STUDENTS					
MEAN	217.6 (1.2)	220.7 (2.0)	221.3 (1.4)	227.1 (1.0)	226.7 (1.0)
ST. DEV.	44.6 (0.8)	39.8 (1.3)	41.1 (0.8)	38.4 (0.5)	38.8 (0.6)
PERCENTILES					
5	141.3 (3.5)	151.2 (6.6)	152.5 (2.5)	159.9 (2.4)	161.0 (3.4)
10	158.5 (2.2)	167.5 (3.1)	166.9 (2.6)	175.8 (2.2)	175.3 (2.2)
25	188.3 (1.4)	195.3 (2.6)	193.2 (1.8)	201.9 (1.2)	200.9 (1.5)
50	219.5 (1.2)	221.4 (3.6)	222.5 (2.0)	229.2 (1.1)	228.5 (1.4)
75	248.6 (1.1)	247.4 (2.4)	250.2 (1.9)	254.0 (1.1)	253.7 (1.5)
90	273.8 (1.6)	270.6 (3.4)	273.3 (1.6)	274.6 (1.9)	275.0 (1.7)
95	288.2 (1.6)	284.4 (3.3)	287.0 (2.6)	287.0 (1.9)	287.7 (1.2)

Table 16-91 (continued)

Percentiles for Science Long-term Trend, Grade 4

	1977	1982	1986	1990	1992
WHITE STUDENTS					
MEAN	229.6 (0.9)	229.0 (1.9)	231.9 (1.2)	237.5 (0.8)	239.1 (1.0)
ST. DEV.	40.0 (0.5)	37.6 (1.3)	39.2 (0.7)	36.3 (0.4)	36.4 (0.5)
PERCENTILES					
5	163.2 (1.3)	167.0 (3.0)	166.5 (2.3)	176.9 (1.4)	178.0 (2.0)
10	177.6 (1.1)	182.2 (3.1)	181.0 (1.5)	189.9 (1.3)	191.7 (1.5)
25	202.4 (1.1)	203.8 (2.6)	205.5 (1.5)	212.6 (0.8)	214.5 (1.3)
50	229.8 (0.9)	228.6 (2.4)	232.5 (1.6)	238.3 (1.0)	240.0 (1.1)
75	256.9 (0.8)	254.9 (2.0)	258.8 (1.4)	262.3 (1.0)	264.2 (1.3)
90	281.1 (1.1)	277.6 (2.8)	281.7 (1.7)	283.5 (1.4)	285.1 (1.6)
95	295.4 (1.9)	290.8 (4.0)	294.9 (2.5)	295.7 (1.3)	297.5 (0.8)
BLACK STUDENTS					
MEAN	174.8 (1.8)	187.0 (3.0)	196.2 (1.9)	196.4 (2.0)	200.3 (2.7)
ST. DEV.	41.4 (1.0)	37.7 (1.9)	38.3 (1.0)	38.6 (1.0)	37.3 (0.7)
PERCENTILES					
5	107.0 (3.5)	123.6 (11.0)	132.8 (3.2)	131.3 (4.2)	138.0 (4.2)
10	122.8 (3.4)	136.7 (8.3)	146.9 (3.5)	145.3 (3.8)	151.6 (4.0)
25	146.6 (2.4)	159.2 (4.9)	169.7 (2.6)	169.8 (2.6)	173.7 (3.5)
50	173.8 (2.5)	188.2 (5.0)	195.9 (2.2)	196.3 (2.5)	201.1 (3.0)
75	202.9 (1.8)	214.4 (3.8)	222.6 (1.5)	224.1 (1.7)	226.3 (3.4)
90	229.2 (2.9)	236.4 (4.7)	246.4 (3.7)	246.8 (2.4)	248.4 (3.0)
95	244.1 (2.9)	246.5 (3.3)	259.5 (3.5)	260.0 (5.4)	260.5 (4.6)
HISPANIC STUDENTS					
MEAN	191.9 (2.7)	189.0 (4.2)	199.4 (3.1)	206.2 (2.2)	204.7 (2.8)
ST. DEV.	41.2 (1.4)	36.6 (2.3)	38.9 (1.6)	37.0 (1.7)	37.3 (1.4)
PERCENTILES					
5	125.2 (7.0)	127.3 (9.6)	134.1 (10.1)	146.2 (5.5)	143.0 (3.0)
10	139.8 (3.3)	141.9 (16.8)	148.1 (5.2)	158.6 (4.3)	156.8 (3.9)
25	163.9 (4.3)	161.9 (7.4)	172.6 (3.4)	180.6 (3.7)	179.1 (3.5)
50	191.4 (3.6)	190.8 (4.8)	199.8 (6.7)	206.2 (3.7)	204.8 (4.1)
75	219.0 (3.2)	215.9 (3.4)	225.6 (4.1)	232.7 (4.1)	230.4 (2.3)
90	245.7 (4.9)	236.2 (5.6)	252.1 (5.4)	252.9 (4.4)	253.7 (5.5)
95	261.3 (6.4)	246.0 (7.6)	264.9 (6.7)	266.8 (6.9)	264.9 (3.5)

Table 16-92

Percentiles for Science Long-term Trend, Grade 8

	1977	1982	1986	1990	1992
TOTAL SAMPLE					
MEAN	247.4 (1.1)	250.1 (1.3)	251.4 (1.4)	255.2 (0.9)	258.0 (0.8)
ST. DEV.	43.5 (0.4)	38.6 (0.5)	36.6 (0.6)	37.6 (0.7)	36.9 (0.5)
PERCENTILES					
5	173.7 (1.7)	185.2 (2.2)	188.9 (2.2)	191.4 (2.0)	193.1 (1.5)
10	190.6 (1.4)	199.6 (1.8)	203.3 (2.0)	205.9 (1.7)	208.9 (1.3)
25	218.4 (1.4)	224.1 (1.1)	227.2 (1.3)	230.0 (1.5)	234.7 (1.3)
50	248.6 (1.2)	250.9 (1.3)	252.1 (1.8)	256.4 (1.2)	260.4 (1.0)
75	277.5 (0.9)	276.7 (1.5)	276.5 (1.5)	281.1 (0.9)	283.8 (1.0)
90	302.4 (0.9)	299.2 (1.6)	298.2 (2.0)	302.4 (1.1)	303.1 (1.2)
95	316.9 (1.5)	312.8 (1.3)	310.3 (1.6)	315.1 (1.9)	314.6 (1.4)
MALE STUDENTS					
MEAN	251.1 (1.3)	255.6 (1.5)	256.1 (1.6)	258.5 (1.1)	260.1 (1.2)
ST. DEV.	43.9 (0.5)	38.7 (0.6)	37.4 (1.0)	38.8 (0.8)	38.0 (0.8)
PERCENTILES					
5	176.7 (1.9)	190.2 (2.6)	192.3 (4.2)	191.9 (2.5)	193.4 (2.7)
10	193.5 (1.6)	204.4 (1.6)	207.2 (2.5)	207.3 (3.4)	209.4 (2.4)
25	221.5 (1.7)	229.5 (1.7)	231.1 (1.6)	232.9 (1.4)	235.8 (1.1)
50	252.4 (1.5)	256.7 (1.5)	256.9 (2.0)	260.3 (1.4)	262.7 (1.5)
75	281.6 (1.2)	282.6 (1.5)	282.4 (1.4)	285.8 (2.2)	287.0 (1.8)
90	306.5 (1.3)	305.0 (1.7)	303.4 (1.6)	307.4 (1.5)	306.4 (1.8)
95	321.2 (1.5)	318.3 (2.3)	316.2 (2.2)	320.2 (1.2)	318.1 (1.6)
FEMALE STUDENTS					
MEAN	243.7 (1.2)	245.0 (1.3)	246.9 (1.5)	251.8 (1.1)	256.0 (1.0)
ST. DEV.	42.8 (0.5)	37.9 (0.7)	35.3 (0.6)	36.1 (0.8)	35.7 (0.8)
PERCENTILES					
5	170.8 (1.6)	180.2 (1.9)	186.3 (2.1)	190.6 (2.1)	192.7 (1.6)
10	187.7 (1.8)	195.5 (2.3)	200.5 (2.9)	204.8 (1.5)	208.4 (1.4)
25	215.5 (1.7)	219.7 (1.4)	223.4 (1.5)	227.8 (1.6)	233.4 (1.3)
50	245.0 (1.2)	246.1 (1.7)	248.0 (1.7)	253.1 (1.2)	258.2 (1.4)
75	273.0 (1.5)	271.0 (1.9)	271.0 (1.8)	276.8 (1.6)	280.7 (1.9)
90	297.7 (1.0)	292.8 (1.5)	291.3 (1.7)	296.8 (1.1)	299.8 (1.1)
95	312.1 (2.2)	305.3 (1.8)	304.0 (3.6)	308.6 (1.4)	311.1 (1.7)

Table 16-92 (continued)

Percentiles for Science Long-term Trend, Grade 8

	1977	1982	1986	1990	1992
WHITE STUDENTS					
MEAN	256.1 (0.8)	257.3 (1.1)	259.2 (1.4)	264.1 (0.9)	267.1 (1.0)
ST. DEV.	39.5 (0.3)	35.7 (0.6)	33.6 (0.8)	33.8 (0.5)	31.8 (0.6)
PERCENTILES					
5	190.8 (0.9)	198.0 (1.7)	203.5 (2.7)	208.6 (1.6)	212.6 (2.2)
10	205.2 (1.2)	210.8 (1.7)	215.8 (1.5)	220.4 (1.2)	225.7 (1.6)
25	229.3 (1.3)	233.2 (1.2)	237.0 (1.9)	241.3 (0.9)	246.1 (1.1)
50	256.3 (0.8)	257.6 (1.3)	259.2 (2.0)	264.5 (1.1)	267.8 (1.1)
75	282.9 (0.7)	281.5 (1.1)	282.3 (1.9)	287.0 (1.7)	289.0 (1.2)
90	306.6 (0.9)	302.7 (1.6)	302.2 (1.9)	307.1 (1.4)	307.1 (1.6)
95	320.8 (1.1)	316.2 (1.7)	313.9 (2.1)	319.4 (1.3)	318.0 (1.4)
BLACK STUDENTS					
MEAN	208.1 (2.4)	217.1 (1.3)	221.6 (2.5)	225.7 (3.1)	224.4 (2.7)
ST. DEV.	39.7 (0.9)	34.6 (1.2)	33.0 (0.9)	34.3 (1.7)	37.1 (1.3)
PERCENTILES					
5	144.3 (3.2)	160.3 (3.1)	167.8 (1.7)	169.7 (5.5)	162.1 (3.7)
10	157.7 (2.4)	173.0 (3.1)	180.1 (2.2)	181.8 (6.1)	177.0 (3.8)
25	180.5 (2.2)	193.7 (2.4)	198.3 (3.0)	202.3 (3.7)	198.9 (3.6)
50	207.4 (2.5)	216.8 (1.3)	221.2 (2.8)	225.7 (3.0)	223.8 (2.4)
75	234.8 (2.6)	240.7 (2.2)	243.5 (3.6)	249.1 (2.6)	251.4 (3.6)
90	259.5 (3.4)	262.2 (3.5)	264.4 (4.9)	269.0 (4.2)	272.0 (2.7)
95	274.6 (2.7)	274.7 (1.9)	276.8 (2.5)	283.2 (3.7)	286.0 (7.6)
HISPANIC STUDENTS					
MEAN	213.4 (1.9)	225.5 (3.9)	226.1 (3.1)	231.6 (2.6)	237.5 (2.6)
ST. DEV.	40.4 (1.2)	36.2 (1.1)	34.2 (1.2)	36.6 (1.0)	34.0 (1.2)
PERCENTILES					
5	147.1 (3.5)	166.3 (4.9)	171.1 (5.6)	173.7 (4.7)	180.3 (3.7)
10	161.4 (3.0)	179.4 (4.1)	181.3 (4.5)	185.3 (4.5)	193.0 (6.4)
25	185.8 (3.5)	200.7 (3.6)	201.6 (5.5)	205.9 (4.1)	215.2 (3.8)
50	213.3 (2.5)	225.9 (4.4)	225.6 (3.8)	230.9 (3.3)	237.9 (4.5)
75	240.3 (3.5)	249.3 (5.1)	249.8 (3.4)	256.4 (5.1)	260.9 (3.4)
90	265.8 (2.0)	271.2 (5.1)	269.9 (3.5)	280.0 (5.9)	281.8 (2.5)
95	282.1 (4.4)	284.8 (6.1)	283.0 (3.8)	294.2 (2.8)	292.1 (4.2)

Table 16-93

Percentiles for Science Long-term Trend, Grade 12

	1977	1982	1986	1990	1992
TOTAL SAMPLE					
MEAN	289.5 (1.0)	283.3 (1.2)	288.5 (1.4)	290.4 (1.1)	294.1 (1.3)
ST. DEV.	45.0 (0.4)	46.7 (0.7)	44.4 (1.0)	46.2 (0.6)	44.7 (0.8)
PERCENTILES					
5	212.6 (1.3)	203.2 (2.2)	211.8 (2.4)	209.9 (2.3)	217.7 (2.1)
10	231.3 (1.4)	221.5 (1.9)	229.5 (2.4)	228.8 (2.0)	234.2 (2.5)
25	260.6 (1.4)	252.5 (2.1)	259.6 (1.9)	260.3 (1.9)	263.6 (2.3)
50	290.8 (1.0)	285.4 (1.0)	290.1 (1.9)	292.2 (1.3)	295.9 (1.5)
75	320.1 (0.9)	315.3 (1.6)	319.4 (1.3)	322.7 (1.4)	326.6 (1.3)
90	346.2 (1.1)	341.5 (1.1)	344.5 (1.9)	348.3 (1.2)	350.3 (1.9)
95	361.5 (1.3)	357.3 (1.4)	359.9 (2.0)	362.9 (1.5)	363.8 (1.2)
MALE STUDENTS					
MEAN	297.0 (1.2)	291.9 (1.4)	294.9 (1.9)	295.6 (1.3)	299.1 (1.7)
ST. DEV.	45.3 (0.6)	47.1 (0.9)	46.6 (1.2)	48.7 (0.9)	46.3 (1.0)
PERCENTILES					
5	219.5 (2.1)	210.3 (2.3)	213.9 (2.8)	210.4 (3.9)	219.0 (3.9)
10	238.2 (1.6)	228.9 (2.7)	231.4 (5.0)	229.5 (2.9)	235.5 (4.2)
25	267.6 (1.5)	261.1 (1.9)	263.5 (3.0)	263.4 (1.3)	267.4 (3.0)
50	298.5 (1.2)	294.3 (1.4)	298.7 (2.8)	297.9 (1.9)	301.3 (2.2)
75	328.1 (1.4)	324.8 (2.0)	327.6 (1.6)	329.9 (1.8)	333.6 (1.4)
90	353.9 (1.4)	350.6 (1.9)	353.4 (2.8)	356.7 (2.3)	357.2 (1.0)
95	368.8 (1.6)	365.3 (1.3)	367.0 (4.6)	372.5 (1.8)	370.4 (1.5)
FEMALE STUDENTS					
MEAN	282.2 (1.1)	275.2 (1.3)	282.3 (1.5)	285.4 (1.6)	289.0 (1.5)
ST. DEV.	43.5 (0.5)	44.8 (0.8)	41.3 (1.1)	43.2 (1.0)	42.3 (1.2)
PERCENTILES					
5	207.5 (1.6)	198.3 (3.6)	209.8 (3.5)	209.2 (3.7)	216.5 (4.2)
10	226.1 (2.1)	215.5 (2.6)	228.1 (2.0)	228.2 (4.5)	232.9 (2.8)
25	254.5 (1.5)	245.7 (2.1)	256.2 (2.0)	257.7 (2.4)	260.3 (2.4)
50	283.8 (1.2)	277.6 (2.0)	283.7 (1.4)	287.7 (2.0)	290.9 (2.1)
75	311.5 (1.1)	306.2 (1.2)	310.8 (1.8)	316.2 (2.3)	319.8 (1.9)
90	336.3 (1.2)	330.1 (1.0)	333.5 (3.0)	339.6 (2.3)	341.4 (1.9)
95	351.2 (1.5)	345.2 (1.5)	348.3 (3.2)	351.5 (1.6)	354.4 (2.2)

Table 16-93 (continued)

Percentiles for Science Long-term Trend, Grade 12

	1977	1982	1986	1990	1992
WHITE STUDENTS					
MEAN	297.7 (0.7)	293.1 (1.0)	297.5 (1.7)	300.9 (1.1)	304.2 (1.3)
ST. DEV.	40.5 (0.3)	41.6 (0.5)	40.6 (1.0)	41.1 (0.6)	40.6 (0.9)
PERCENTILES					
5	231.1 (0.9)	223.0 (1.7)	228.3 (2.9)	232.8 (2.3)	234.3 (3.9)
10	246.0 (0.7)	239.1 (1.5)	244.5 (3.1)	249.0 (2.0)	251.3 (2.5)
25	270.3 (0.8)	265.5 (1.5)	271.0 (2.0)	273.4 (1.5)	276.8 (2.2)
50	297.5 (0.7)	293.6 (1.0)	298.7 (1.7)	301.2 (1.2)	306.0 (1.5)
75	325.0 (0.9)	321.2 (1.6)	324.9 (1.3)	329.0 (1.6)	333.0 (1.7)
90	349.9 (1.0)	346.0 (1.3)	348.9 (3.0)	352.3 (1.3)	355.1 (1.5)
95	364.6 (1.4)	360.8 (1.3)	363.5 (2.8)	367.3 (2.0)	368.5 (0.9)
BLACK STUDENTS					
MEAN	240.2 (1.5)	234.7 (1.7)	252.8 (2.9)	253.0 (4.5)	256.2 (3.2)
ST. DEV.	41.6 (0.9)	41.8 (1.3)	40.4 (2.2)	44.7 (2.4)	39.4 (1.4)
PERCENTILES					
5	172.4 (1.5)	166.0 (3.1)	189.3 (4.8)	182.0 (10.1)	191.8 (4.0)
10	187.3 (1.9)	180.6 (3.5)	201.6 (4.9)	196.6 (3.1)	206.6 (4.1)
25	212.1 (1.4)	206.4 (3.2)	225.0 (4.2)	220.5 (4.3)	230.1 (1.7)
50	240.4 (1.8)	234.7 (3.0)	251.9 (5.9)	251.6 (3.0)	255.4 (3.2)
75	267.9 (2.0)	262.7 (2.2)	279.5 (3.4)	282.9 (6.0)	282.4 (5.9)
90	293.4 (2.6)	288.8 (3.9)	306.0 (4.2)	313.6 (11.3)	308.2 (10.3)
95	309.6 (2.6)	305.4 (1.6)	322.8 (5.8)	329.3 (10.2)	324.8 (8.7)
HISPANIC STUDENTS					
MEAN	262.3 (2.2)	248.7 (2.3)	259.3 (3.8)	261.5 (4.4)	270.2 (5.6)
ST. DEV.	41.8 (1.5)	43.4 (2.3)	39.3 (1.7)	44.1 (2.6)	41.6 (2.0)
PERCENTILES					
5	193.7 (5.2)	178.0 (6.1)	194.4 (9.3)	188.7 (6.2)	196.6 (10.5)
10	208.4 (4.0)	194.2 (7.2)	209.2 (3.8)	203.9 (11.1)	215.4 (14.6)
25	234.3 (3.9)	218.8 (3.3)	232.0 (5.6)	230.6 (3.6)	241.6 (8.6)
50	262.4 (2.4)	248.0 (2.5)	258.9 (5.8)	260.5 (5.7)	272.7 (11.0)
75	289.5 (5.1)	278.4 (3.4)	285.8 (3.6)	292.6 (10.6)	297.9 (2.8)
90	316.9 (4.4)	302.1 (3.4)	309.9 (7.6)	317.4 (5.1)	322.8 (6.7)
95	331.3 (4.4)	320.8 (11.0)	324.4 (6.3)	329.5 (9.1)	339.1 (6.0)

Table 16-94

Weighted Means, Standard Deviations, and Percentiles of Writing Distributions for Writing Main Samples, by Subgroup, Grade 4

	MEAN	STD DEV	5TH	10TH	25TH	50TH	75TH	90TH	95TH
-- TOTAL --	221.8(1.2)	40.8(0.6)	153.1(2.0)	169.3(1.9)	195.0(1.4)	222.4(1.3)	249.8(1.1)	273.4(1.5)	287.1(1.1)
GENDER									
MALE	215.4(1.2)	40.0(0.6)	148.3(1.8)	163.6(1.9)	188.9(1.6)	216.3(1.9)	243.2(1.2)	266.2(2.0)	279.9(1.4)
FEMALE	228.0(1.6)	40.6(0.8)	159.4(3.6)	175.5(2.8)	201.6(2.1)	228.6(1.4)	255.8(1.6)	279.8(1.4)	292.9(2.6)
RACE/ETHNICITY									
WHITE	229.4(1.3)	38.3(0.7)	165.9(1.7)	180.0(1.5)	203.9(1.8)	229.8(1.9)	255.6(1.2)	278.6(1.2)	290.8(1.8)
BLACK	195.3(2.3)	38.0(1.1)	132.1(3.2)	146.3(2.9)	169.9(3.1)	196.0(2.1)	221.1(3.2)	244.2(3.0)	257.1(4.0)
HISPANIC	207.8(2.0)	40.8(1.8)	140.6(3.7)	155.7(2.8)	181.7(3.1)	208.8(2.9)	235.7(2.9)	259.1(3.3)	273.0(2.6)
ASIAN/PACIFIC AMERICAN	233.4(3.9)	40.0(2.7)	165.7(14.4)	178.2(10.2)	206.1(6.6)	235.4(5.7)	261.5(5.2)	281.0(6.3)	296.5(8.5)
AMER IND/ALASKAN NATV	215.0(3.5)	39.5(4.0)	154.2(11.5)	166.2(8.3)	188.4(6.4)	213.3(5.1)	241.2(8.6)	265.4(12.0)	279.0(10.9)
PARENTS' EDUCATION LEVEL									
LESS THAN H.S.	208.0(3.5)	37.8(2.5)	145.1(11.3)	158.3(6.0)	185.0(8.1)	209.1(4.5)	233.0(10.3)	255.6(3.8)	268.6(6.7)
GRADUATED H.S.	213.5(1.6)	38.6(1.9)	148.2(4.4)	163.2(4.5)	187.7(2.4)	214.1(2.6)	239.7(3.0)	262.9(4.2)	276.8(5.8)
SOME EDUC AFTER H.S.	228.0(2.5)	39.5(1.3)	161.8(5.1)	177.1(3.2)	201.9(1.7)	228.6(3.4)	255.3(3.0)	277.5(4.5)	291.8(13.1)
GRADUATED COLLEGE	231.6(1.2)	39.9(1.0)	164.6(4.7)	179.5(2.1)	205.1(2.2)	233.1(1.8)	259.0(2.7)	282.1(1.3)	294.4(2.6)
UNKNOWN	213.4(1.6)	39.8(0.7)	146.7(2.2)	161.6(1.5)	186.8(2.2)	214.2(1.7)	240.6(1.9)	263.6(2.7)	278.4(2.1)

Table 16-95

Weighted Means, Standard Deviations, and Percentiles of Writing Distributions for Writing Main Samples, by Subgroup, Grade 8

	MEAN	STD DEV	5TH	10TH	25TH	50TH	75TH	90TH	95TH
-- TOTAL --	261.6(1.0)	37.9(0.5)	197.6(1.5)	212.6(1.5)	236.4(1.1)	262.2(1.0)	287.6(1.1)	310.0(1.4)	322.8(1.9)
GENDER									
MALE	251.5(1.1)	37.0(0.6)	189.7(1.3)	203.9(1.2)	226.9(1.1)	251.9(1.5)	276.7(1.2)	298.7(1.3)	312.6(2.3)
FEMALE	271.8(1.2)	35.9(0.7)	211.4(2.5)	225.4(2.7)	248.1(1.2)	272.5(1.3)	296.4(1.3)	317.2(1.8)	329.8(2.1)
RACE/ETHNICITY									
WHITE	268.0(1.1)	36.2(0.4)	207.7(2.2)	221.3(1.4)	244.0(1.5)	268.4(1.2)	292.8(1.2)	314.4(1.6)	326.8(1.7)
BLACK	241.7(2.2)	35.6(1.1)	182.7(2.3)	195.3(2.6)	217.5(3.5)	242.3(2.9)	266.2(3.1)	287.4(3.2)	299.5(2.5)
HISPANIC	247.9(1.7)	37.0(1.0)	186.3(3.9)	200.3(2.3)	223.1(2.0)	248.1(2.2)	272.8(2.4)	295.5(1.6)	308.7(3.5)
ASIAN/PACIFIC AMERICAN	267.4(4.6)	39.7(2.7)	199.7(6.8)	213.9(4.2)	240.3(5.9)	268.2(5.1)	295.0(5.4)	318.0(6.6)	330.3(8.4)
AMER IND/ALASKAN NATV	248.5(4.3)	39.3(2.8)	177.1(7.3)	193.7(36.3)	221.4(5.4)	253.9(6.5)	276.4(4.1)	297.0(7.2)	306.9(5.2)
PARENTS' EDUCATION LEVEL									
LESS THAN H.S.	250.5(1.7)	35.9(1.5)	190.1(4.7)	203.6(4.2)	226.6(2.1)	251.2(3.4)	274.0(1.9)	295.3(3.5)	308.8(3.9)
GRADUATED H.S.	254.4(2.0)	37.0(0.9)	191.9(1.8)	206.4(4.3)	229.3(2.1)	254.8(2.1)	280.2(2.7)	301.5(3.1)	315.1(3.9)
SOME EDUC AFTER H.S.	264.6(1.0)	35.4(0.6)	205.7(1.9)	218.8(1.6)	242.0(1.6)	264.8(1.2)	288.9(1.0)	309.8(1.9)	321.8(3.8)
GRADUATED COLLEGE	270.8(1.4)	37.0(0.6)	208.6(1.9)	223.0(1.8)	246.2(1.8)	271.7(1.7)	296.5(1.3)	317.7(1.4)	329.7(2.6)
UNKNOWN	240.8(2.2)	36.9(1.8)	180.4(5.2)	193.5(4.8)	216.2(2.7)	240.9(2.7)	265.7(3.5)	287.9(4.1)	301.2(4.2)

Table 16-96

Weighted Means, Standard Deviations, and Percentiles of Writing Distributions for Writing Main Samples, by Subgroup, Grade 12

	MEAN	STD DEV	5TH	10TH	25TH	50TH	75TH	90TH	95TH
-- TOTAL --	286.1(1.0)	40.5(0.4)	218.1(2.0)	233.3(1.4)	259.1(1.5)	287.2(1.1)	313.8(1.0)	337.5(1.0)	350.8(1.5)
GENDER									
MALE	275.3(1.2)	40.1(0.7)	208.0(3.1)	223.7(1.6)	248.0(1.2)	275.8(1.3)	303.3(1.1)	326.3(1.4)	340.0(1.7)
FEMALE	296.3(1.2)	38.2(0.6)	232.4(2.4)	246.5(2.1)	271.0(1.9)	297.1(1.0)	322.6(1.2)	344.7(1.2)	357.9(1.9)
RACE/ETHNICITY									
WHITE	291.0(1.1)	39.4(0.5)	224.6(2.0)	239.7(1.5)	265.0(1.4)	292.1(1.4)	317.8(1.2)	340.9(1.4)	354.0(1.2)
BLACK	267.6(1.8)	39.0(1.0)	204.4(2.6)	219.1(3.0)	241.3(1.8)	267.8(2.5)	293.0(2.7)	317.6(3.0)	332.7(3.8)
HISPANIC	276.9(1.8)	40.0(1.9)	209.4(5.9)	223.9(4.2)	250.2(1.9)	278.1(3.4)	305.3(2.8)	327.6(2.9)	340.2(4.6)
ASIAN/PACIFIC AMERICAN	291.6(2.9)	41.2(1.8)	220.9(7.0)	236.8(3.5)	264.6(4.8)	293.5(3.5)	319.7(2.2)	342.9(2.5)	358.0(6.5)
AMER IND/ALASKAN NATV	269.5(8.2)	42.2(4.1)	199.2(14.8)	214.1(9.6)	240.5(9.7)	269.9(11.8)	298.0(7.0)	323.8(11.7)	338.0(10.9)
PARENTS' EDUCATION LEVEL									
LESS THAN H.S.	271.0(1.9)	38.8(1.3)	208.9(4.7)	221.3(2.4)	243.6(2.5)	271.4(2.2)	298.7(2.8)	320.5(3.3)	332.5(3.6)
GRADUATED H.S.	274.8(1.3)	39.2(0.8)	208.8(3.1)	224.0(1.8)	248.0(1.5)	275.1(2.1)	301.6(1.7)	324.5(2.3)	339.3(1.9)
SOME EDUC AFTER H.S.	287.1(1.7)	37.9(0.8)	223.6(2.7)	237.2(2.3)	261.8(1.8)	287.9(2.4)	313.2(1.9)	335.1(2.1)	347.6(1.6)
GRADUATED COLLEGE	296.4(1.2)	39.3(0.5)	229.3(1.2)	245.2(1.6)	270.5(1.5)	297.6(1.5)	323.3(1.5)	345.9(1.6)	359.0(1.2)
UNKNOWN	250.3(3.1)	40.1(2.2)	179.6(4.0)	197.4(10.1)	225.6(5.2)	251.2(3.6)	278.1(2.0)	301.6(6.0)	312.5(12.8)

Table 16-97

Percentiles for Writing Long-term Trend, Grade 4

	1984	1988	1990	1992
TOTAL SAMPLE				
MEAN	203.8 (1.5)	205.7 (1.6)	201.7 (1.5)	207.1 (1.5)
ST. DEV.	36.5 (1.4)	42.0 (0.8)	41.7 (0.8)	38.2 (0.6)
PERCENTILES				
5	143.5 (3.3)	135.4 (3.3)	130.8 (1.8)	142.2 (1.9)
10	156.8 (2.7)	151.3 (1.9)	147.2 (2.3)	157.3 (1.3)
25	179.2 (2.1)	177.0 (1.6)	174.2 (2.1)	181.9 (1.4)
50	204.2 (2.9)	207.0 (2.5)	202.9 (1.7)	208.1 (2.2)
75	228.7 (1.5)	235.0 (1.8)	230.8 (2.3)	233.2 (1.6)
90	250.5 (2.2)	259.1 (2.1)	254.6 (2.1)	255.8 (1.8)
95	262.9 (2.7)	273.7 (2.6)	268.3 (2.7)	268.8 (2.9)
MALE STUDENTS				
MEAN	200.5 (2.8)	199.0 (2.3)	195.0 (1.9)	198.3 (1.7)
ST. DEV.	36.4 (1.6)	41.5 (1.0)	41.6 (1.1)	37.4 (0.8)
PERCENTILES				
5	140.8 (7.2)	129.1 (4.3)	124.6 (4.0)	135.0 (2.3)
10	153.7 (4.3)	145.5 (3.2)	140.4 (3.1)	149.4 (2.2)
25	176.7 (2.9)	170.7 (2.5)	167.8 (2.8)	173.7 (1.8)
50	200.0 (3.1)	199.2 (3.8)	195.9 (1.9)	199.6 (1.9)
75	225.5 (3.1)	228.0 (2.9)	223.8 (1.6)	223.5 (2.3)
90	246.3 (2.8)	251.5 (2.7)	247.7 (1.9)	245.6 (3.0)
95	260.2 (3.4)	265.7 (3.5)	262.2 (2.5)	258.0 (3.0)
FEMALE STUDENTS				
MEAN	207.6 (3.1)	212.6 (2.0)	208.7 (2.2)	216.1 (1.7)
ST. DEV.	36.2 (2.3)	41.4 (1.2)	40.6 (1.1)	36.9 (0.9)
PERCENTILES				
5	146.7 (6.1)	143.3 (2.2)	139.8 (3.6)	153.0 (3.1)
10	161.1 (5.5)	158.2 (3.4)	155.4 (3.6)	167.6 (1.5)
25	183.5 (3.4)	184.9 (4.5)	181.8 (2.7)	191.7 (2.5)
50	208.7 (3.0)	213.5 (2.5)	209.8 (2.9)	217.7 (1.7)
75	232.3 (3.4)	241.3 (1.9)	237.0 (2.7)	241.6 (1.5)
90	253.7 (7.0)	265.4 (3.8)	260.2 (2.0)	262.9 (1.9)
95	265.1 (5.5)	279.7 (2.5)	273.1 (2.7)	275.4 (2.9)
WHITE STUDENTS				
MEAN	210.7 (1.9)	214.9 (1.9)	211.0 (2.0)	216.7 (1.7)
ST. DEV.	34.1 (1.1)	38.7 (1.2)	38.2 (0.8)	34.7 (0.7)
PERCENTILES				
5	154.9 (3.6)	151.0 (4.3)	146.4 (3.0)	159.0 (2.3)
10	166.6 (3.3)	164.6 (3.1)	161.6 (3.5)	172.1 (1.7)
25	187.6 (1.5)	188.5 (1.8)	185.6 (2.9)	193.9 (1.9)
50	211.1 (2.2)	215.5 (1.8)	211.4 (1.9)	217.3 (2.3)
75	233.2 (1.4)	241.6 (2.1)	237.3 (2.2)	240.4 (1.6)
90	255.0 (4.5)	264.6 (2.6)	260.0 (1.6)	260.8 (2.2)
95	266.1 (3.5)	278.4 (3.3)	272.5 (2.4)	273.1 (2.8)
BLACK STUDENTS				
MEAN	181.6 (5.0)	173.3 (4.7)	171.4 (5.4)	175.0 (3.8)
ST. DEV.	34.8 (3.3)	39.4 (2.3)	40.3 (1.6)	33.9 (1.6)
PERCENTILES				
5	123.9 (6.5)	108.9 (7.6)	104.8 (5.2)	117.2 (5.7)
10	134.9 (4.5)	122.3 (5.9)	119.5 (6.8)	130.5 (4.2)
25	159.5 (4.9)	147.7 (5.5)	144.0 (4.9)	152.5 (4.5)
50	181.5 (5.5)	172.6 (4.6)	172.0 (6.0)	175.7 (4.2)
75	205.2 (13.6)	199.7 (4.6)	198.5 (7.0)	198.2 (3.3)
90	227.9 (10.2)	224.4 (6.0)	222.7 (6.5)	217.9 (3.6)
95	239.6 (12.0)	237.5 (5.8)	239.0 (23.2)	229.2 (6.4)
HISPANIC STUDENTS				
MEAN	188.5 (5.8)	190.3 (3.5)	184.1 (4.1)	189.4 (3.6)
ST. DEV.	37.1 (5.3)	39.6 (1.7)	38.8 (1.9)	35.3 (1.6)
PERCENTILES				
5	130.0 (13.1)	125.3 (6.0)	120.3 (7.5)	132.5 (5.1)
10	141.2 (12.3)	139.4 (6.1)	134.8 (7.6)	143.7 (5.1)
25	161.9 (8.7)	163.2 (5.1)	158.7 (7.1)	166.4 (3.4)
50	188.2 (6.4)	190.8 (5.1)	183.7 (5.4)	188.8 (4.3)
75	214.2 (8.0)	217.5 (4.5)	210.1 (3.8)	213.2 (5.1)
90	234.5 (6.5)	240.7 (4.3)	234.2 (4.8)	234.5 (5.5)
95	247.3 (7.5)	256.5 (4.8)	247.8 (5.8)	247.3 (3.5)

Table 16-98
Percentiles for Writing Long-term Trend, Grade 8

	1984	1988	1990	1992
TOTAL SAMPLE				
MEAN	266.7 (2.0)	263.7 (1.3)	256.6 (1.2)	274.4 (1.3)
ST. DEV.	29.5 (1.4)	32.4 (0.6)	37.5 (0.7)	36.3 (0.9)
PERCENTILES				
5	215.7 (4.3)	208.6 (2.9)	194.6 (2.3)	213.7 (2.0)
10	227.0 (3.7)	221.9 (2.6)	207.9 (1.9)	227.2 (2.9)
25	247.3 (3.1)	242.5 (1.9)	231.4 (1.4)	249.9 (1.5)
50	268.3 (2.6)	264.1 (1.4)	256.7 (1.6)	275.0 (1.6)
75	287.5 (2.1)	285.7 (1.2)	281.9 (1.7)	299.7 (1.7)
90	303.7 (1.8)	305.0 (1.4)	304.4 (1.3)	320.4 (1.7)
95	312.9 (2.0)	315.7 (1.8)	318.4 (1.8)	332.4 (1.6)
MALE STUDENTS				
MEAN	257.5 (2.3)	253.7 (1.5)	245.6 (1.5)	263.6 (1.9)
ST. DEV.	28.9 (1.7)	31.3 (0.9)	36.1 (1.0)	34.8 (1.0)
PERCENTILES				
5	208.0 (4.4)	201.0 (3.7)	187.0 (1.7)	205.8 (3.3)
10	219.7 (4.1)	213.0 (2.5)	199.3 (1.8)	218.7 (2.2)
25	237.8 (5.7)	233.1 (1.6)	220.7 (2.2)	240.3 (2.6)
50	258.6 (3.2)	254.4 (1.4)	245.6 (1.6)	263.7 (2.3)
75	278.0 (2.7)	275.0 (1.3)	270.0 (1.8)	288.2 (2.1)
90	293.9 (3.0)	293.6 (2.1)	291.9 (2.8)	307.9 (1.8)
95	302.9 (2.4)	304.0 (3.5)	304.9 (2.5)	319.5 (2.3)
FEMALE STUDENTS				
MEAN	276.2 (2.4)	273.5 (1.7)	267.9 (1.3)	285.0 (1.3)
ST. DEV.	27.0 (1.4)	30.5 (0.6)	35.4 (1.1)	34.6 (1.0)
PERCENTILES				
5	229.3 (4.9)	222.6 (3.3)	210.1 (2.6)	227.0 (4.1)
10	241.5 (4.4)	234.3 (1.5)	223.1 (2.3)	239.9 (3.2)
25	258.7 (2.7)	253.3 (2.4)	244.2 (1.5)	261.7 (1.6)
50	277.1 (3.5)	274.1 (2.3)	267.8 (1.2)	285.7 (2.5)
75	295.7 (3.0)	294.4 (1.5)	291.7 (1.7)	308.9 (1.5)
90	309.7 (2.4)	312.3 (1.2)	313.6 (1.8)	328.9 (1.3)
95	317.6 (3.5)	322.5 (2.0)	327.3 (2.2)	340.1 (3.0)

	1984	1988	1990	1992
WHITE STUDENTS				
MEAN	271.7 (2.1)	269.1 (1.3)	262.1 (1.6)	279.2 (1.3)
ST. DEV.	27.5 (1.1)	31.1 (0.6)	36.4 (0.7)	35.1 (0.8)
PERCENTILES				
5	224.4 (5.7)	216.5 (2.3)	202.1 (2.3)	220.4 (3.4)
10	235.1 (3.2)	229.0 (1.4)	215.3 (2.9)	234.0 (2.6)
25	253.4 (2.3)	248.5 (1.8)	237.3 (1.9)	255.5 (1.7)
50	272.7 (2.7)	269.9 (1.5)	262.5 (1.6)	280.1 (1.7)
75	291.0 (2.3)	290.5 (1.3)	286.9 (1.8)	303.5 (1.9)
90	306.3 (2.0)	308.7 (1.7)	308.5 (1.8)	323.7 (2.4)
95	315.0 (2.5)	319.1 (1.3)	322.5 (1.8)	335.4 (2.5)
BLACK STUDENTS				
MEAN	247.1 (5.7)	246.0 (3.5)	239.0 (2.3)	258.1 (4.0)
ST. DEV.	27.4 (1.7)	31.0 (1.1)	35.6 (1.9)	36.5 (2.2)
PERCENTILES				
5	200.8 (7.2)	193.6 (5.2)	181.6 (5.7)	200.1 (8.1)
10	212.1 (6.8)	205.2 (6.1)	192.9 (4.0)	212.5 (7.4)
25	228.5 (6.9)	226.0 (3.4)	215.6 (4.4)	232.2 (4.7)
50	247.6 (6.7)	246.6 (4.3)	239.8 (2.4)	257.0 (3.6)
75	265.3 (3.7)	266.3 (3.3)	263.3 (2.7)	282.1 (4.6)
90	281.3 (5.4)	285.2 (5.7)	284.5 (1.6)	306.3 (3.3)
95	291.5 (3.6)	295.7 (4.4)	297.2 (3.1)	318.9 (4.7)
HISPANIC STUDENTS				
MEAN	246.9 (6.4)	250.4 (2.5)	245.7 (2.8)	265.0 (2.2)
ST. DEV.	30.2 (2.9)	31.0 (1.6)	36.1 (1.7)	35.3 (2.5)
PERCENTILES				
5	197.2 (13.9)	199.2 (5.9)	186.7 (5.4)	203.4 (6.4)
10	206.6 (11.7)	209.8 (7.1)	198.8 (3.7)	218.7 (5.1)
25	225.0 (7.6)	230.5 (5.3)	220.5 (3.6)	242.0 (5.0)
50	247.2 (7.0)	251.2 (2.9)	246.5 (4.7)	265.1 (2.1)
75	268.4 (4.2)	270.9 (2.4)	269.7 (5.0)	288.2 (4.3)
90	286.1 (5.8)	289.5 (2.4)	292.3 (3.3)	310.2 (3.3)
95	298.1 (7.7)	300.6 (5.4)	304.6 (3.3)	323.8 (3.5)

Table 16-99
Percentiles for Writing Long-term Trend, Grade 12

	1984	1988	1990	1992
TOTAL SAMPLE				
MEAN	289.7 (1.6)	291.3 (1.3)	287.1 (1.0)	287.3 (1.4)
ST. DEV.	31.8 (0.6)	27.9 (1.1)	36.5 (1.1)	32.0 (0.7)
PERCENTILES				
5	235.7 (2.3)	244.4 (4.5)	226.7 (2.7)	233.2 (3.3)
10	248.9 (1.7)	255.2 (2.4)	240.3 (2.4)	245.7 (2.1)
25	268.9 (1.7)	273.1 (1.9)	262.3 (1.8)	266.1 (1.7)
50	290.9 (1.9)	292.4 (1.3)	287.7 (1.3)	288.2 (1.2)
75	312.1 (1.7)	310.6 (1.2)	311.9 (1.4)	309.7 (1.7)
90	329.5 (2.6)	326.3 (1.4)	334.0 (1.8)	327.6 (1.5)
95	340.0 (2.9)	334.8 (1.6)	346.6 (2.7)	338.1 (2.3)
MALE STUDENTS				
MEAN	281.1 (1.4)	282.2 (2.0)	276.4 (1.6)	279.4 (1.2)
ST. DEV.	31.5 (0.8)	26.8 (1.0)	35.2 (0.8)	31.2 (1.1)
PERCENTILES				
5	227.2 (3.5)	236.7 (4.2)	218.2 (2.8)	226.9 (3.2)
10	240.7 (2.4)	248.2 (3.5)	231.3 (2.6)	239.0 (2.0)
25	260.3 (1.4)	264.9 (2.4)	252.8 (2.0)	258.6 (1.7)
50	281.5 (1.9)	282.9 (2.0)	276.2 (1.8)	280.5 (1.2)
75	303.0 (2.0)	300.6 (1.7)	300.9 (1.6)	301.3 (1.7)
90	320.8 (3.2)	316.0 (1.9)	321.4 (1.6)	318.5 (1.6)
95	331.4 (2.6)	324.5 (2.6)	333.1 (1.8)	328.6 (2.2)
FEMALE STUDENTS				
MEAN	298.6 (2.5)	299.3 (1.2)	298.2 (1.5)	296.4 (2.0)
ST. DEV.	29.5 (1.1)	26.4 (1.3)	34.4 (1.6)	30.4 (0.6)
PERCENTILES				
5	249.1 (4.4)	254.2 (3.8)	241.2 (4.0)	244.0 (1.9)
10	261.4 (4.0)	265.4 (2.7)	253.0 (3.0)	256.6 (3.1)
25	279.3 (3.1)	282.0 (1.8)	274.9 (1.6)	276.1 (2.2)
50	299.5 (2.8)	300.7 (1.7)	298.4 (1.6)	297.4 (2.3)
75	318.9 (2.8)	317.8 (1.8)	321.8 (1.5)	317.6 (2.1)
90	335.0 (4.1)	331.7 (1.7)	342.6 (2.6)	334.9 (1.8)
95	344.8 (4.9)	340.5 (1.7)	354.3 (4.1)	344.3 (2.8)

	1984	1988	1990	1992
WHITE STUDENTS				
MEAN	296.8 (1.8)	296.2 (1.3)	292.8 (1.2)	294.1 (1.2)
ST. DEV.	28.5 (1.0)	26.2 (1.2)	35.2 (1.2)	29.6 (0.9)
PERCENTILES				
5	249.2 (3.8)	252.2 (3.7)	234.7 (3.4)	243.8 (1.2)
10	260.0 (2.2)	262.7 (2.8)	247.2 (2.7)	255.6 (2.0)
25	277.0 (1.4)	278.9 (1.7)	268.9 (1.6)	274.8 (1.1)
50	297.6 (2.0)	297.1 (1.2)	293.7 (1.1)	295.0 (1.3)
75	316.4 (2.6)	314.5 (1.3)	317.0 (1.3)	314.2 (1.8)
90	333.0 (2.6)	329.3 (1.3)	338.2 (2.2)	331.4 (1.4)
95	343.0 (4.0)	337.7 (2.8)	349.6 (2.8)	341.0 (2.5)
BLACK STUDENTS				
MEAN	270.3 (3.6)	275.2 (2.9)	268.2 (2.3)	263.2 (3.2)
ST. DEV.	29.6 (2.4)	26.0 (1.4)	34.3 (1.1)	28.9 (1.4)
PERCENTILES				
5	221.8 (8.5)	231.7 (6.2)	213.1 (5.4)	216.4 (5.0)
10	231.8 (6.8)	242.8 (2.9)	224.7 (3.0)	226.3 (3.2)
25	251.6 (5.5)	257.5 (3.3)	244.9 (2.3)	244.7 (5.8)
50	270.0 (3.3)	275.7 (3.0)	268.4 (3.2)	263.8 (3.9)
75	289.6 (4.6)	293.8 (3.0)	291.4 (4.0)	282.6 (4.3)
90	308.9 (7.8)	308.9 (4.9)	311.0 (2.3)	300.5 (4.1)
95	317.6 (5.0)	318.1 (3.8)	324.0 (4.4)	309.3 (5.1)
HISPANIC STUDENTS				
MEAN	259.1 (6.6)	273.8 (4.4)	276.9 (2.6)	273.6 (3.8)
ST. DEV.	30.8 (2.0)	28.1 (1.7)	36.2 (2.3)	31.3 (1.5)
PERCENTILES				
5	207.5 (10.8)	228.3 (6.2)	217.4 (9.7)	220.4 (10.6)
10	216.5 (6.5)	235.9 (8.4)	231.5 (4.1)	234.0 (2.7)
25	238.2 (7.5)	256.1 (6.2)	252.6 (4.0)	252.4 (6.0)
50	260.3 (8.2)	274.1 (3.9)	275.4 (3.2)	275.3 (4.6)
75	281.1 (7.1)	294.3 (8.7)	300.8 (6.0)	294.1 (4.9)
90	297.3 (9.8)	309.0 (5.6)	324.2 (3.9)	313.7 (3.5)
95	305.7 (16.8)	316.3 (5.4)	337.7 (15.8)	323.7 (3.1)

Table 16-100

Weighted Response Percentages and Proficiency Means (with Standard Errors) for Reading Main Samples, by Gender, Grade 4

	N	WTD %	(CV)	MALE	FEMALE	MISSING
-- TOTAL --	6314	100.0(0.0)	[0%]	51.1(0.6) 213.9(1.2)	48.9(0.6) 222.0(1.0)	0.0 *****
GENDER						
MALE	3171	51.1(0.6)	[1%]	100.0(0.0) 213.9(1.2)	0.0(0.0) *****(0.0)	0.0 *****
FEMALE	3143	48.9(0.6)	[1%]	0.0(0.0) *****(0.0)	100.0(0.0) 222.0(1.0)	0.0 *****
RACE/ETHNICITY						
WHITE	3917	70.6(0.2)	[1%]	51.3(0.7) 222.2(1.4)	48.7(0.7) 229.7(1.4)	0.0 *****
BLACK	1013	15.8(0.1)	[1%]	46.5(1.7) 186.3(2.3)	53.5(1.7) 199.4(2.2)	0.0 *****
HISPANIC	1044	9.4(0.1)	[2%]	54.6(1.9) 197.5(3.2)	45.4(1.9) 207.3(2.4)	0.0 *****
ASIAN/PACIFIC AMERICAN	201	2.3(0.3)	[12%]	60.7(3.2) 213.6(4.0)	39.3(3.2) 218.6(5.2)	0.0 *****
AMER IND/ALASKAN NATV	117	1.7(0.2)	[15%]	55.4(6.3) 198.5(5.1)	44.6(6.3) 218.6(5.9)	0.0 *****
PARENTS' EDUCATION LEVEL						
LESS THAN H.S.	291	4.1(0.4)	[9%]	43.4(3.9) 190.3(4.2)	56.6(3.9) 205.3(2.9)	0.0 *****
GRADUATED H.S.	769	12.4(0.6)	[5%]	50.9(1.8) 208.0(2.3)	49.1(1.8) 218.8(2.6)	0.0 *****
SOME EDUC AFTER H.S.	529	8.6(0.5)	[6%]	57.5(2.4) 220.0(2.7)	42.5(2.4) 229.3(2.7)	0.0 *****
GRADUATED COLLEGE	2481	39.0(1.1)	[3%]	52.6(1.2) 223.1(1.8)	47.4(1.2) 230.7(1.4)	0.0 *****
UNKNOWN	2228	35.7(1.0)	[3%]	48.7(1.1) 206.2(1.7)	51.3(1.1) 215.5(1.5)	0.0 *****

BEST COPY AVAILABLE

Table 16-101

Weighted Response Percentages and Proficiency Means (with Standard Errors) for Reading Main Samples, by Derived Race/ethnicity, Grade 4

	N	WEIGHTED PCT [CV]	WHITE	BLACK	HISPANIC	ASIAN	AMER IND	UNCLASSI	MISSING
-- TOTAL --	6314	100.0(0.0) [0%]	70.6(0.2) 225.9(1.2)	15.8(0.1) 193.3(1.7)	9.4(0.1) 202.0(2.2)	2.3(0.3) 215.5(3.3)	1.7(0.2) 207.5(4.7)	0.2(0.1) 208.2(5.7)	0.0 *****
GENDER									
MALE	3171	51.1(0.6) [1%]	70.8(0.6) 222.2(1.4)	14.4(0.5) 186.3(2.3)	10.1(0.4) 197.5(3.2)	2.7(0.4) 213.6(4.0)	1.8(0.3) 198.5(5.1)	0.2(0.1) 216.8(7.1)	0.0 *****
FEMALE	3143	48.9(0.6) [1%]	70.3(0.6) 229.7(1.4)	17.3(0.5) 199.4(2.2)	8.7(0.4) 207.3(2.4)	1.8(0.3) 218.6(5.2)	1.5(0.3) 218.6(5.9)	0.3(0.1) 203.0(10.7)	0.0 *****
RACE/ETHNICITY									
WHITE	3917	70.6(0.2) [1%]	100.0(0.0) 225.9(1.2)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
BLACK	1013	15.8(0.1) [1%]	0.0(0.0) *****(0.0)	100.0(0.0) 193.3(1.7)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
HISPANIC	1044	9.4(0.1) [2%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 202.0(2.2)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
ASIAN/PACIFIC AMERICAN	201	2.3(0.3) [12%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 215.5(3.3)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
AMER IND/ALASKAN NATV	117	1.7(0.2) [15%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 207.5(4.7)	0.0(0.0) *****(0.0)	0.0 *****
PARENTS' EDUCATION LEVEL									
LESS THAN H.S.	291	4.1(0.4) [9%]	58.1(3.8) 203.9(4.1)	17.4(2.8) 186.1(6.8)	19.3(2.2) 193.8(4.3)	3.0(1.2) 197.2(10.1)	2.0(0.9) 208.8(25.5)	0.2(0.2) 237.0(****)	0.0 *****
GRADUATED H.S.	769	12.4(0.6) [5%]	70.0(1.7) 221.1(1.9)	19.8(1.8) 191.0(2.9)	7.5(0.9) 207.2(4.6)	0.9(0.3) 201.3(15.6)	1.8(0.5) 187.8(17.1)	0.1(0.1) 216.8(7.6)	0.0 *****
SOME EDUC AFTER H.S.	529	8.6(0.5) [6%]	78.1(2.0) 230.1(2.5)	10.5(1.5) 193.4(4.3)	8.2(1.0) 209.6(4.7)	1.6(0.5) 211.1(11.6)	1.2(0.5) 201.6(15.6)	0.3(0.2) 238.1(10.1)	0.0 *****
GRADUATED COLLEGE	2481	39.0(1.1) [3%]	72.9(1.0) 234.8(1.6)	15.6(0.9) 198.0(2.5)	7.1(0.5) 210.1(3.6)	2.3(0.3) 227.2(4.6)	1.9(0.3) 215.4(7.3)	0.3(0.1) 226.4(10.9)	0.0 *****
UNKNOWN	2228	35.7(1.0) [3%]	68.0(1.2) 218.4(1.5)	15.8(0.9) 190.8(2.7)	11.8(0.6) 196.1(2.7)	2.8(0.4) 210.1(3.7)	1.5(0.3) 206.8(5.6)	0.1(0.0) 190.6(17.2)	0.0 *****

Table 16-102

**Weighted Response Percentages and Proficiency Means (with Standard Errors)
for Reading Main Samples, by Parents' Education Level, Grade 4**

	N	WEIGHTED PCT [CV]	NOT HS	GRAD HS	POST HS	GRAD COL	UNKNOWN	MISSING
-- TOTAL --	6298	100.0(0.0) [0%]	4.1(0.4) 198.8(2.7)	12.4(0.6) 213.3(1.7)	8.6(0.5) 224.0(2.2)	39.0(1.1) 226.7(1.4)	35.8(1.0) 211.0(1.2)	0.2 155.0
GENDER								
MALE	3163	51.1(0.6) [1%]	3.5(0.4) 190.3(4.2)	12.4(0.7) 208.0(2.3)	9.7(0.7) 220.0(2.7)	40.2(1.2) 223.1(1.8)	34.2(1.2) 206.2(1.7)	0.2 159.9
FEMALE	3135	48.9(0.6) [1%]	4.8(0.6) 205.3(2.9)	12.5(0.8) 218.8(2.6)	7.5(0.7) 229.3(2.7)	37.8(1.4) 230.7(1.4)	37.5(1.3) 215.5(1.5)	0.2 149.4
RACE/ETHNICITY								
WHITE	3913	70.7(0.2) [1%]	3.4(0.4) 203.9(4.1)	12.3(0.7) 221.1(1.9)	9.5(0.7) 230.1(2.5)	40.3(1.5) 234.8(1.6)	34.4(1.3) 218.4(1.5)	0.1 169.9
BLACK	1008	15.8(0.1) [1%]	4.5(0.8) 186.1(6.8)	15.6(1.7) 191.0(2.9)	5.7(0.9) 193.4(4.3)	38.4(1.9) 198.0(2.5)	35.7(2.1) 190.8(2.7)	0.4 148.1
HISPANIC	1043	9.4(0.1) [2%]	8.4(1.1) 193.8(4.3)	9.9(1.2) 207.2(4.6)	7.5(0.9) 209.6(4.7)	29.3(1.8) 210.1(3.6)	44.9(2.1) 196.1(2.7)	0.2 109.0
ASIAN/PACIFIC AMERICAN	200	2.3(0.3) [12%]	5.5(1.9) 197.2(10.1)	4.7(1.7) 201.3(15.6)	6.2(1.6) 211.1(11.6)	38.8(3.9) 227.2(4.6)	44.8(3.5) 210.1(3.7)	0.5 189.2
AMER IND/ALASKAN NATV	116	1.7(0.2) [15%]	5.0(2.1) 208.8(25.5)	13.1(3.4) 187.8(17.1)	6.0(2.4) 201.6(15.6)	44.6(4.9) 215.4(7.3)	31.2(4.1) 206.8(5.6)	0.6 131.6
PARENTS' EDUCATION LEVEL								
LESS THAN H.S.	291	4.1(0.4) [9%]	100.0(0.0) 198.8(2.7)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
GRADUATED H.S.	769	12.4(0.6) [5%]	0.0(0.0) *****(0.0)	100.0(0.0) 213.3(1.7)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
SOME EDUC AFTER H.S.	529	8.6(0.5) [6%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 224.0(2.2)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
GRADUATED COLLEGE	2481	39.0(1.1) [3%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 226.7(1.4)	0.0(0.0) *****(0.0)	0.0 *****
UNKNOWN	2228	35.8(1.0) [3%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 211.0(1.2)	0.0 *****

Table 16-103

Weighted Response Percentages and Proficiency Means (with Standard Errors) for Reading Main Samples, by Gender, Grade 8

	N	WTD %	(CV)	MALE	FEMALE	MISSING
-- TOTAL --	9464	100.0(0.0)	[0%]	50.8(0.7) 254.0(1.1)	49.2(0.7) 266.7(1.0)	0.0 *****
GENDER						
MALE	4689	50.8(0.7)	[1%]	100.0(0.0) 254.0(1.1)	0.0(0.0) *****(0.0)	0.0 *****
FEMALE	4775	49.2(0.7)	[1%]	0.0(0.0) *****(0.0)	100.0(0.0) 266.7(1.0)	0.0 *****
RACE/ETHNICITY						
WHITE	5994	69.9(0.2)	[0%]	51.0(0.8) 261.2(1.3)	49.0(0.8) 274.3(1.3)	0.0 *****
BLACK	1568	15.5(0.2)	[1%]	48.8(1.3) 231.6(2.0)	51.2(1.3) 244.1(1.7)	0.0 *****
HISPANIC	1402	10.2(0.2)	[2%]	51.6(1.4) 235.7(1.8)	48.4(1.4) 247.8(2.0)	0.0 *****
ASIAN/PACIFIC AMERICAN	352	3.0(0.2)	[6%]	51.7(3.2) 262.4(3.8)	48.3(3.2) 279.1(3.9)	0.0 *****
AMER IND/ALASKAN NATV	106	1.1(0.2)	[14%]	54.0(5.5) 242.9(4.2)	46.0(5.5) 260.9(6.3)	0.0 *****
PARENTS' EDUCATION LEVEL						
LESS THAN H.S.	746	7.9(0.5)	[6%]	42.5(1.9) 234.5(3.1)	57.5(1.9) 248.7(1.8)	0.0 *****
GRADUATED H.S.	2137	23.7(0.8)	[3%]	49.1(1.1) 244.6(1.4)	50.9(1.1) 257.1(1.7)	0.0 *****
SOME EDUC AFTER H.S.	1806	19.1(0.5)	[2%]	46.7(1.4) 259.1(1.7)	53.3(1.4) 271.2(1.3)	0.0 *****
GRADUATED COLLEGE	3946	41.3(1.2)	[3%]	53.8(0.9) 264.5(1.2)	46.2(0.9) 278.6(1.2)	0.0 *****
UNKNOWN	805	7.8(0.4)	[6%]	58.0(1.8) 231.6(2.3)	42.0(1.8) 245.9(2.6)	0.0 *****

Table 16-104

Weighted Response Percentages and Proficiency Means (with Standard Errors)
for Reading Main Samples, by Derived Race/ethnicity, Grade 8

	N	WEIGHTED PCT [CV]	WHITE	BLACK	HISPANIC	ASIAN	AMER IND	UNCLASSI	MISSING
-- TOTAL --	9464	100.0(0.0) [0%]	69.9(0.2) 267.6(1.2)	15.5(0.2) 238.0(1.6)	10.2(0.2) 241.6(1.4)	3.0(0.2) 270.5(3.1)	1.1(0.2) 251.2(3.7)	0.4(0.1) 261.6(5.8)	0.0 *****
GENDER									
MALE	4689	50.8(0.7) [1%]	70.2(0.6) 261.2(1.3)	14.9(0.4) 231.6(2.0)	10.3(0.4) 235.7(1.8)	3.0(0.2) 262.4(3.8)	1.2(0.2) 242.9(4.2)	0.4(0.1) 259.9(11.4)	0.0 *****
FEMALE	4775	49.2(0.7) [1%]	69.6(0.5) 274.3(1.3)	16.1(0.5) 244.1(1.7)	10.0(0.3) 247.8(2.0)	2.9(0.3) 279.1(3.9)	1.0(0.2) 260.9(6.3)	0.4(0.1) 263.3(8.2)	0.0 *****
RACE/ETHNICITY									
WHITE	5994	69.9(0.2) [0%]	100.0(0.0) 267.6(1.2)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
BLACK	1568	15.5(0.2) [1%]	0.0(0.0) *****(0.0)	100.0(0.0) 238.0(1.6)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
HISPANIC	1402	10.2(0.2) [2%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 241.6(1.4)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
ASIAN/PACIFIC AMERICAN	352	3.0(0.2) [6%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 270.5(3.1)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
AMER IND/ALASKAN NATV	106	1.1(0.2) [14%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 251.2(3.7)	0.0(0.0) *****(0.0)	0.0 *****
PARENTS' EDUCATION LEVEL									
LESS THAN H.S.	746	7.9(0.5) [6%]	56.7(2.4) 249.8(2.3)	14.3(1.6) 223.4(5.6)	25.4(1.9) 236.0(3.0)	1.9(0.5) 262.7(7.6)	1.2(0.4) 240.4(9.4)	0.5(0.3) 255.6(4.6)	0.0 *****
GRADUATED H.S.	2137	23.7(0.8) [3%]	69.0(1.3) 257.7(1.7)	18.7(1.1) 232.8(2.0)	9.6(0.7) 237.3(2.3)	1.1(0.2) 257.0(8.7)	1.4(0.3) 246.2(5.5)	0.3(0.1) 266.4(6.8)	0.0 *****
SOME EDUC AFTER H.S.	1806	19.1(0.5) [2%]	73.4(0.9) 270.8(1.4)	15.4(0.6) 247.2(2.2)	8.3(0.6) 253.5(2.5)	1.3(0.3) 270.1(6.1)	1.1(0.3) 261.7(8.5)	0.5(0.2) 254.5(19.6)	0.0 *****
GRADUATED COLLEGE	3946	41.3(1.2) [3%]	75.5(0.9) 276.7(1.2)	13.2(0.6) 244.1(1.9)	6.0(0.4) 255.1(2.5)	4.1(0.6) 280.2(3.1)	0.9(0.2) 253.9(6.1)	0.3(0.1) 279.1(8.0)	0.0 *****
UNKNOWN	805	7.8(0.4) [6%]	48.9(2.5) 245.8(2.1)	18.7(1.7) 225.8(4.7)	22.8(1.5) 224.4(2.4)	7.7(1.5) 252.0(5.7)	1.2(0.5) 246.3(9.0)	0.6(0.3) 235.1(13.5)	0.0 *****

Table 16-105

Weighted Response Percentages and Proficiency Means (with Standard Errors) for Reading Main Samples, by Parents' Education Level, Grade 8

	N	WEIGHTED PCT [CV]	NOT HS	GRAD HS	POST HS	GRAD COL	UNKNOWN	MISSING
-- TOTAL --	9440	100.0(0.0) [0%]	8.0(0.5) 242.7(1.5)	23.8(0.8) 250.9(1.4)	19.1(0.5) 265.6(1.1)	41.4(1.2) 271.0(1.0)	7.8(0.4) 237.6(2.0)	0.2 215.2
GENDER								
MALE	4676	50.8(0.7) [1%]	6.7(0.5) 234.5(3.1)	23.0(0.9) 244.6(1.4)	17.6(0.5) 259.1(1.7)	43.8(1.2) 264.5(1.2)	8.9(0.5) 231.6(2.3)	0.2 207.1
FEMALE	4764	49.2(0.7) [1%]	9.3(0.7) 248.7(1.8)	24.6(0.9) 257.1(1.7)	20.7(0.7) 271.2(1.3)	38.8(1.3) 278.6(1.2)	6.7(0.5) 245.9(2.6)	0.2 223.4
RACE/ETHNICITY								
WHITE	5988	69.9(0.2) [0%]	6.4(0.5) 249.8(2.3)	23.4(1.0) 257.7(1.7)	20.0(0.6) 270.8(1.4)	44.6(1.5) 276.7(1.2)	5.5(0.5) 245.8(2.1)	0.1 234.1
BLACK	1561	15.5(0.2) [1%]	7.4(0.8) 223.4(5.6)	28.7(1.6) 232.8(2.0)	19.1(0.9) 247.2(2.2)	35.4(1.6) 244.1(1.9)	9.5(0.9) 225.8(4.7)	0.3 175.5
HISPANIC	1392	10.1(0.2) [2%]	19.9(2.0) 236.0(3.0)	22.5(1.4) 237.3(2.3)	15.7(1.1) 253.5(2.5)	24.3(1.5) 255.1(2.5)	17.6(0.9) 224.4(2.4)	0.6 230.3
ASIAN/PACIFIC AMERICAN	351	3.0(0.2) [6%]	5.0(1.4) 262.7(7.6)	9.1(2.1) 257.0(8.7)	8.0(1.7) 270.1(6.1)	57.5(7.0) 280.2(3.1)	20.3(4.8) 252.0(5.7)	0.2 173.7
AMER IND/ALASKAN NATV	106	1.1(0.2) [14%]	9.0(2.5) 240.4(9.4)	29.2(4.9) 246.2(5.5)	18.2(3.1) 261.7(8.5)	34.8(6.2) 253.9(6.1)	8.8(3.4) 246.3(9.0)	0.0 *****
PARENTS' EDUCATION LEVEL								
LESS THAN H.S.	746	8.0(0.5) [6%]	100.0(0.0) 242.7(1.5)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
GRADUATED H.S.	2137	23.8(0.8) [3%]	0.0(0.0) *****(0.0)	100.0(0.0) 250.9(1.4)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
SOME EDUC AFTER H.S.	1806	19.1(0.5) [2%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 265.6(1.1)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
GRADUATED COLLEGE	3946	41.4(1.2) [3%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 271.0(1.0)	0.0(0.0) *****(0.0)	0.0 *****
UNKNOWN	805	7.8(0.4) [6%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 237.6(2.0)	0.0 *****

Table 16-106

Weighted Response Percentages and Proficiency Means (with Standard Errors) for Reading Main Samples, by Gender, Grade 12

	N	WTD %	(CV)	MALE	FEMALE	MISSING
-- TOTAL --	9856	100.0(0.0)	[1%]	48.6(0.6) 285.8(0.7)	51.4(0.6) 296.0(0.7)	0.0 *****
GENDER						
MALE	4713	48.6(0.6)	[2%]	100.0(0.0) 285.8(0.7)	0.0(0.0) *****(0.0)	0.0 *****
FEMALE	5143	51.4(0.6)	[1%]	0.0(0.0) *****(0.0)	100.0(0.0) 296.0(0.7)	0.0 *****
RACE/ETHNICITY						
WHITE	6728	71.6(0.4)	[1%]	49.4(0.7) 291.4(0.8)	50.6(0.7) 302.4(0.7)	0.0 *****
BLACK	1603	15.0(0.4)	[3%]	44.3(1.5) 266.1(1.4)	55.7(1.5) 276.3(1.9)	0.0 *****
HISPANIC	1005	8.8(0.4)	[5%]	47.9(2.0) 272.8(2.8)	52.1(2.0) 281.2(2.8)	0.0 *****
ASIAN/PACIFIC AMERICAN	426	3.7(0.2)	[4%]	48.3(3.1) 283.4(3.6)	51.7(3.1) 297.2(3.4)	0.0 *****
AMER IND/ALASKAN NATV	59	0.5(0.1)	[20%]	68.0(7.0) 266.3(6.8)	32.0(7.0) 283.4(6.8)	0.0 *****
PARENTS' EDUCATION LEVEL						
LESS THAN H.S.	771	7.5(0.4)	[6%]	39.9(2.2) 267.5(2.2)	60.1(2.2) 277.5(1.9)	0.0 *****
GRADUATED H.S.	2063	21.9(0.5)	[3%]	49.8(1.2) 276.9(1.1)	50.2(1.2) 286.0(1.0)	0.0 *****
SOME EDUC AFTER H.S.	2551	26.7(0.6)	[2%]	46.6(1.3) 286.8(1.2)	53.4(1.3) 298.0(1.0)	0.0 *****
GRADUATED COLLEGE	4190	41.3(0.9)	[2%]	50.6(1.0) 294.9(1.0)	49.4(1.0) 306.2(1.1)	0.0 *****
UNKNOWN	251	2.4(0.2)	[7%]	52.2(3.6) 251.7(3.4)	47.8(3.6) 262.5(4.0)	0.0 *****

Table 16-107

Weighted Response Percentages and Proficiency Means (with Standard Errors) for Reading Main Samples, by Derived Race/ethnicity, Grade 12

	N	WEIGHTED PCT [CV]	WHITE	BLACK	HISPANIC	ASIAN	AMER IND	UNCLASSI	MISSING
-- TOTAL --	9856	100.0(0.0) [1%]	71.6(0.4) 297.0(0.6)	15.0(0.4) 271.8(1.5)	8.8(0.4) 277.2(2.4)	3.7(0.2) 290.6(3.2)	0.5(0.1) 271.8(5.3)	0.4(0.1) 285.4(6.8)	0.0 *****
GENDER									
MALE	4713	48.6(0.6) [2%]	72.7(0.8) 291.4(0.8)	13.7(0.6) 266.1(1.4)	8.7(0.6) 272.8(2.8)	3.7(0.3) 283.4(3.6)	0.7(0.1) 266.3(6.8)	0.6(0.2) 284.6(6.8)	0.0 *****
FEMALE	5143	51.4(0.6) [1%]	70.6(0.6) 302.4(0.7)	16.3(0.5) 276.3(1.9)	8.9(0.4) 281.2(2.8)	3.7(0.3) 297.2(3.4)	0.3(0.1) 283.4(6.8)	0.2(0.1) 287.6(13.5)	0.0 *****
RACE/ETHNICITY									
WHITE	6728	71.6(0.4) [1%]	100.0(0.0) 297.0(0.6)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
BLACK	1603	15.0(0.4) [3%]	0.0(0.0) *****(0.0)	100.0(0.0) 271.8(1.5)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
HISPANIC	1005	8.8(0.4) [5%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 277.2(2.4)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
ASIAN/PACIFIC AMERICAN	426	3.7(0.2) [4%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 290.6(3.2)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
AMER IND/ALASKAN NATV	59	0.5(0.1) [20%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 271.8(5.3)	0.0(0.0) *****(0.0)	0.0 *****
PARENTS' EDUCATION LEVEL									
LESS THAN H.S.	771	7.5(0.4) [6%]	47.8(2.9) 280.8(2.0)	17.0(1.9) 259.1(3.1)	30.5(3.3) 269.8(2.9)	3.5(0.7) 274.7(8.1)	1.1(0.5) 275.5(14.8)	0.0(0.0) *****(0.0)	0.0 *****
GRADUATED H.S.	2063	21.9(0.5) [3%]	71.5(1.2) 286.8(1.0)	16.9(1.0) 262.0(2.0)	8.9(0.7) 274.5(2.5)	2.1(0.3) 288.0(6.5)	0.4(0.1) 253.9(18.3)	0.2(0.1) 298.3(11.9)	0.0 *****
SOME EDUC AFTER H.S.	2551	26.7(0.6) [2%]	73.1(1.0) 297.5(0.8)	17.3(0.8) 277.7(2.1)	7.0(0.6) 283.4(3.9)	1.8(0.2) 288.0(4.9)	0.5(0.2) 275.8(11.3)	0.3(0.2) 285.8(22.4)	0.0 *****
GRADUATED COLLEGE	4190	41.3(0.9) [2%]	76.9(0.8) 304.9(0.9)	11.7(0.6) 279.8(1.7)	5.4(0.4) 286.5(3.6)	5.1(0.3) 299.1(3.4)	0.4(0.1) 277.0(8.6)	0.5(0.2) 293.2(9.0)	0.0 *****
UNKNOWN	251	2.4(0.2) [7%]	39.9(3.6) 261.9(3.8)	24.5(3.4) 249.8(5.9)	18.9(3.2) 255.6(7.1)	14.7(3.8) 259.5(5.8)	0.4(0.4) 262.3(26.2)	1.7(1.3) 230.8(14.2)	0.0 *****

Table 16-108

Weighted Response Percentages and Proficiency Means (with Standard Errors) for Reading Main Samples, by Parents' Education Level, Grade 12

	N	WEIGHTED PCT [CV]	NOT HS	GRAD HS	POST HS	GRAD COL	UNKNOWN	MISSING
-- TOTAL --	9826	100.0(0.0) [1%]	7.6(0.4) 273.5(1.5)	22.0(0.5) 281.4(0.8)	26.7(0.6) 292.7(0.8)	41.4(0.9) 300.5(0.8)	2.4(0.2) 256.9(2.8)	0.3 245.8
GENDER								
MALE	4694	48.6(0.6) [2%]	6.2(0.4) 267.5(2.2)	22.5(0.7) 276.9(1.1)	25.6(0.8) 286.8(1.2)	43.1(1.1) 294.9(1.0)	2.5(0.2) 251.7(3.4)	0.3 249.4
FEMALE	5132	51.4(0.6) [1%]	8.8(0.6) 277.5(1.9)	21.5(0.7) 286.0(1.0)	27.8(0.8) 298.0(1.0)	39.8(1.2) 306.2(1.1)	2.2(0.2) 262.5(4.0)	0.2 239.1
RACE/ETHNICITY								
WHITE	6708	71.6(0.4) [1%]	5.0(0.4) 280.8(2.0)	21.9(0.7) 286.8(1.0)	27.3(0.7) 297.5(0.8)	44.4(1.2) 304.9(0.9)	1.3(0.1) 261.9(3.8)	0.3 248.6
BLACK	1598	15.0(0.3) [3%]	8.6(0.9) 259.1(3.1)	24.7(1.2) 262.0(2.0)	30.7(1.4) 277.7(2.1)	32.2(1.7) 279.8(1.7)	3.8(0.6) 249.8(5.9)	0.3 237.3
HISPANIC	1002	8.8(0.4) [5%]	26.2(2.7) 269.8(2.9)	22.3(2.0) 274.5(2.5)	21.2(1.4) 283.4(3.9)	25.3(2.7) 286.5(3.6)	5.1(0.8) 255.6(7.1)	0.1 253.6
ASIAN/PACIFIC AMERICAN	424	3.7(0.2) [4%]	7.2(1.2) 274.7(8.1)	12.4(1.9) 288.0(6.5)	13.3(1.7) 288.0(4.9)	57.7(3.6) 299.1(3.4)	9.4(2.4) 259.5(5.8)	0.4 226.7
AMER IND/ALASKAN NATV	59	0.5(0.1) [20%]	17.9(6.4) 275.5(14.8)	18.9(6.0) 253.9(18.3)	27.7(7.4) 275.8(11.3)	33.8(7.1) 277.0(8.6)	1.8(1.9) 262.3(26.2)	0.0 *****
PARENTS' EDUCATION LEVEL								
LESS THAN H.S.	771	7.6(0.4) [6%]	100.0(0.0) 273.5(1.5)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
GRADUATED H.S.	2063	22.0(0.5) [3%]	0.0(0.0) *****(0.0)	100.0(0.0) 281.4(0.8)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
SOME EDUC AFTER H.S.	2551	26.7(0.6) [2%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 292.7(0.8)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
GRADUATED COLLEGE	4190	41.4(0.9) [2%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 300.5(0.8)	0.0(0.0) *****(0.0)	0.0 *****
UNKNOWN	251	2.4(0.2) [7%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 256.9(2.8)	0.0 *****

Table 16-109
Proficiencies for Reading Long-term Trend Samples, Grade 4

	1971	1975	1980	1984	1988	1990	1992
-- TOTAL --	100.0(0.0) 207.6(1.0)	100.0(0.0) 210.0(0.7)	100.0(0.0) 215.0(1.0)■*	100.0(0.0) 210.9(0.7)■	100.0(0.0) 211.8(1.1)■	100.0(0.0) 209.2(1.2)	100.0(0.0) 210.5(0.9)•
GENDER							
MALE	49.8(0.4) 201.2(1.1)*	50.0(0.4) 204.3(0.8)	50.0(0.4) 210.0(1.1)■	49.9(0.5) 207.5(0.8)■	50.3(0.8) 207.5(1.4)■	50.8(0.9) 204.0(1.7)	50.8(0.7) 205.9(1.3)•
FEMALE	50.2(0.4) 213.9(1.0)	50.0(0.4) 215.8(0.8)	50.0(0.4) 220.1(1.1)■*	50.1(0.5) 214.2(0.8)	49.7(0.8) 216.3(1.3)	49.2(0.9) 214.5(1.2)	49.2(0.7) 215.4(0.9)
RACE/ETHNICITY							
WHITE	84.3(1.4)* 214.0(0.9)*	80.0(1.2)* 216.6(0.7)	79.0(1.3)■* 221.3(0.8)■*	74.9(1.2)■ 218.2(0.8)■	75.2(1.0)■ 217.7(1.4)	74.0(1.0)■ 217.0(1.3)	74.1(0.9)• 217.9(1.0)•
BLACK	13.5(1.3) 170.1(1.7)*	13.4(0.8) 181.2(1.2)■	14.0(1.0) 189.3(1.8)■	15.5(0.5) 185.7(1.1)■	15.9(0.7) 188.5(2.4)■	16.1(0.6) 181.8(2.9)■	15.7(0.4) 184.5(2.2)•
HISPANIC	*****(0.0) *****(0.0)	4.8(0.8) 182.7(2.2)	5.7(0.8) 190.2(2.3)	7.3(1.4) 187.2(2.1)	6.2(1.0) 193.7(3.5)■	6.2(0.6) 189.4(2.3)	6.8(0.8) 191.7(3.1)•
OTHER	2.2(0.6) 193.5(3.8)	1.8(0.5)* 207.8(4.1)	1.3(0.3)* 218.5(3.8)■	2.3(0.2)* 223.8(2.5)■*	2.7(0.4) 228.4(5.4)■*	3.6(0.7) 205.5(4.4)	3.4(0.3) 207.5(4.0)•
PARENTS' EDUCATION LEVEL							
LESS THAN H.S.	10.0(0.4)* 188.6(1.5)	9.9(0.4)* 189.9(1.3)	6.5(0.5)■ 194.3(1.6)■	5.5(0.2)■ 195.1(1.4)■	4.6(0.6)■ 192.5(4.9)	5.0(0.5)■ 192.6(3.2)	5.0(0.4)• 194.9(4.5)
GRADUATED H.S.	22.3(0.5)* 207.8(1.2)	23.8(0.4)* 211.3(0.9)	25.3(0.8)■* 213.0(1.3)■*	19.3(0.6)■* 208.9(1.0)	15.9(0.6)■ 210.8(2.2)	17.1(0.8)■ 209.1(1.8)	15.4(0.8)• 207.4(1.5)
POST H.S.	33.0(0.9)* 223.9(1.1)	34.2(0.7)* 221.5(0.9)	40.1(1.5)■ 226.0(1.1)*	36.2(1.0)* 222.9(0.9)	45.1(1.4)■ 220.0(1.7)	42.3(1.3)■ 217.7(2.0)■	44.3(0.9)• 219.5(1.4)•
UNKNOWN	34.7(0.7) 197.4(1.0)*	32.1(0.8) 203.1(0.8)■-	28.1(1.0)■* 206.1(1.0)■	37.6(0.9) 204.4(0.7)■	34.2(1.3) 204.4(1.5)■	35.5(1.1) 201.4(1.5)	34.4(1.3) 204.1(1.2)•

* Indicates a significant difference between 1992 and a previous year, at the 95% certainty level.

■ Indicates a significant difference between the earliest year with available data and a subsequent year, at the 95% certainty level. These notations (*,■) indicate statistical significance from a multiple comparison procedure based on the number of assessments with valid data.

• Indicates a significant difference between 1992 and the earliest year with availiable data, at the 95% certainty level. This notation indicates statistical significance without application of a multiple comparison procedure.

572

573

Table 16-110

Proficiencies for Reading Long-term Trend Samples, Grade 8

	1971	1975	1980	1984	1988	1990	1992
-- TOTAL --	100.0(0.0) 255.2(0.9)*	100.0(0.0) 255.9(0.8)*	100.0(0.0) 258.5(0.9)	100.0(0.0) 257.1(0.5)	100.0(0.0) 257.5(1.0)	100.0(0.0) 256.8(0.8)	100.0(0.0) 259.8(1.2)•
GENDER							
MALE	50.0(0.4) 249.6(1.0)	49.9(0.5) 249.6(0.8)	49.4(0.3) 254.3(1.1)■	51.0(0.5) 252.6(0.6)■	49.5(0.8) 251.8(1.3)	50.2(0.9) 250.5(1.1)	49.2(0.9) 254.1(1.7)•
FEMALE	50.0(0.4) 260.8(0.9)*	50.1(0.5) 262.3(0.9)	50.6(0.3) 262.6(0.9)	49.0(0.5) 261.7(0.6)	50.5(0.8) 263.0(1.0)	49.8(0.9) 263.1(1.1)	50.8(0.9) 265.3(1.2)•
RACE/ETHNICITY							
WHITE	84.2(1.4)* 260.9(0.7)*	80.9(1.2)* 262.1(0.7)*	79.7(1.8)* 264.4(0.7)■	76.8(0.6)■* 262.6(0.6)*	76.4(0.7)■* 261.3(1.1)*	73.5(0.8)■ 262.3(0.9)*	73.0(0.7)• 266.4(1.2)•
BLACK	14.5(1.4) 222.4(1.2)*	12.7(0.9)* 225.7(1.2)*	13.5(1.3) 232.8(1.5)■	14.1(0.2)* 236.3(1.0)■	15.0(0.3) 242.9(2.4)■	15.4(0.2) 241.5(2.2)■	16.1(0.4) 237.6(2.3)•
HISPANIC	*****(0.0) *****(0.0)	4.9(0.8)* 232.5(3.0)	5.5(1.0) 237.2(2.0)	6.9(0.7) 239.6(1.7)	6.1(0.6) 240.1(3.5)	8.2(0.5)■ 237.8(2.3)	7.3(0.4)• 239.2(3.5)
OTHER	1.2(0.3)* 251.3(3.0)*	1.6(0.5)* 255.6(3.4)*	1.3(0.3)* 253.7(6.4)	1.9(0.2)* 260.0(2.8)	2.5(0.3)■ 269.3(4.2)■	3.0(0.8) 252.7(5.3)	3.5(0.3)• 268.5(3.4)•
PARENTS' EDUCATION LEVEL							
LESS THAN H.S.	16.4(0.6)* 238.4(1.3)	14.0(0.6)■* 238.7(1.2)	10.2(0.6)■* 238.5(1.1)	8.6(0.4)■* 240.0(0.9)	7.9(0.6)■ 246.5(2.1)■	7.9(0.6)■ 240.8(1.8)	6.0(0.5)• 239.2(2.6)
GRADUATED H.S.	31.6(0.7) 255.5(0.8)	33.2(0.6)* 254.6(0.7)	30.7(0.7) 253.5(0.9)	35.3(1.1)■* 253.4(0.7)	30.9(1.0) 252.7(1.2)	30.8(1.2) 251.4(0.9)■	28.1(1.2)• 252.1(1.7)
POST H.S.	38.1(1.1)* 270.2(0.8)	40.0(0.9)* 269.8(0.8)	49.1(1.3)■* 270.9(0.8)	45.1(1.1)■* 267.6(0.7)	51.6(1.5)■ 265.3(1.4)■	50.2(1.5)■* 266.9(1.0)■	56.4(1.6)• 269.9(1.4)
UNKNOWN	14.0(0.8)* 233.1(1.0)	12.7(0.6)* 234.8(1.1)	10.0(0.7)■ 233.3(1.7)	9.8(0.4)■ 236.5(1.3)	9.2(0.7)■ 240.4(3.0)	10.7(0.6)■ 237.7(1.9)	9.1(0.5)• 236.2(2.6)

* Indicates a significant difference between 1992 and a previous year, at the 95% certainty level.

■ Indicates a significant difference between the earliest year with available data and a subsequent year, at the 95% certainty level. These notations (*,■) indicate statistical significance from a multiple comparison procedure based on the number of assessments with valid data.

• Indicates a significant difference between 1992 and the earliest year with availiable data, at the 95% certainty level. This notation indicates statistical significance without application of a multiple comparison procedure.

Table 16-111
Proficiencies for Reading Long-term Trend Samples, Grade 12

	1971	1975	1980	1984	1988	1990	1992
-- TOTAL --	100.0(0.0) 285.2(1.2)*	100.0(0.0) 285.6(0.8)*	100.0(0.0) 285.5(1.2)*	100.0(0.0) 288.8(0.6)■	100.0(0.0) 290.1(1.0)■	100.0(0.0) 290.2(1.1)■	100.0(0.0) 289.7(1.1)•
GENDER							
MALE	49.2(0.5)* 278.9(1.2)*	48.8(0.5)* 279.7(1.0)	50.5(0.6) 281.8(1.3)	51.1(0.7) 283.8(0.6)■	47.7(1.3)* 286.0(1.5)■	50.3(0.6) 284.0(1.6)	52.1(0.9)• 284.2(1.6)•
FEMALE	50.8(0.5)* 291.3(1.3)	51.2(0.5)* 291.2(1.0)*	49.5(0.6) 289.2(1.2)*	48.9(0.7) 293.9(0.8)	52.1(1.4) 293.8(1.5)	49.7(0.6) 296.5(1.2)■	47.9(0.9)• 295.7(1.1)•
RACE/ETHNICITY							
WHITE	87.1(1.3)* 291.4(1.0)*	84.3(1.0)* 293.0(0.6)*	83.1(1.6)* 292.8(0.9)*	77.4(0.6)■* 295.2(0.7)■	76.7(0.6)■ 294.7(1.2)	73.5(0.5)■ 296.6(1.2)■	74.7(0.6)• 297.4(1.4)•
BLACK	11.4(1.2)* 238.7(1.7)*	11.0(0.8)* 240.6(2.0)*	12.0(1.4) 243.1(1.8)*	14.1(0.2) 264.3(1.0)■	15.2(0.3)■ 274.4(2.4)■*	15.9(0.3)■* 267.3(2.3)■	14.7(0.3)• 260.6(2.1)•
HISPANIC	*****(0.0) *****(0.0)	3.4(0.6)* 252.4(3.6)*	3.9(0.6)* 261.4(2.7)	6.6(0.7)■ 268.1(2.2)■	5.8(0.5)■ 270.8(4.3)■	7.0(0.4)■ 274.8(3.6)■	7.5(0.6)• 271.2(3.7)•
OTHER	1.5(0.4)* 275.9(4.8)	1.4(0.4)* 274.3(4.4)	1.1(0.2)* 280.4(3.0)	1.9(0.1)* 284.7(3.1)	2.3(0.3) 290.0(5.3)	3.6(0.5)■ 290.1(3.5)	2.9(0.3)• 287.0(5.8)
PARENTS' EDUCATION LEVEL							
LESS THAN H.S.	19.8(0.8)* 261.3(1.5)	16.0(0.6)■* 262.5(1.3)	12.8(0.7)■* 262.1(1.5)	11.5(0.6)■* 269.4(1.1)■	8.9(0.8)■ 267.4(2.0)	8.8(0.6)■ 269.7(2.8)■	8.1(0.8)• 270.8(3.9)•
GRADUATED H.S.	31.1(0.8) 283.0(1.2)	33.6(0.5)* 281.4(1.1)	32.3(0.9)* 277.5(1.0)■	34.8(1.1)■* 281.2(0.7)	30.2(1.2) 282.0(1.3)	29.8(1.0) 282.9(1.4)	28.3(0.9)• 280.5(1.6)
POST H.S.	41.9(1.3)* 302.2(1.0)	46.4(0.8)■* 300.6(0.7)	51.3(1.3)■* 298.9(1.0)	49.5(1.2)■* 301.2(0.7)	58.2(1.6)■ 299.5(1.3)	57.9(1.3)■ 299.9(1.1)	60.5(1.4)• 298.6(1.4)•
UNKNOWN	7.2(0.8)* 261.1(5.0)	4.0(0.2)■* 239.8(2.8)■	3.6(0.4)■ 249.8(3.5)	3.3(0.2)■ 256.5(2.0)	2.4(0.3)■ 254.7(6.2)	3.2(0.3)■ 245.9(5.7)	2.9(0.3)• 254.7(5.9)

* Indicates a significant difference between 1992 and a previous year, at the 95% certainty level.

■ Indicates a significant difference between the earliest year with available data and a subsequent year, at the 95% certainty level. These notations (*,■) indicate statistical significance from a multiple comparison procedure based on the number of assessments with valid data.

• Indicates a significant difference between 1992 and the earliest year with availiable data, at the 95% certainty level. This notation indicates statistical significance without application of a multiple comparison procedure.

Table 16-112

Weighted Response Percentages and Proficiency Means (with Standard Errors) for Mathematics Main Samples, by Gender, Grade 4

	N	WEIGHTED PCT [CV]	MALE	FEMALE	MISSING
-- TOTAL --	7176	100.0(0.0) [0%]	50.2(0.6) 219.7(0.8)	49.8(0.6) 217.3(1.0)	0.0 *****
GENDER					
MALE	3553	50.2(0.6) [1%]	100.0(0.0) 219.7(0.8)	0.0(0.0) *****(0.0)	0.0 *****
FEMALE	3623	49.8(0.6) [1%]	0.0(0.0) *****(0.0)	100.0(0.0) 217.3(1.0)	0.0 *****
RACE/ETHNICITY					
WHITE	4381	70.4(0.2) [0%]	50.1(0.7) 228.4(1.0)	49.9(0.7) 225.1(1.3)	0.0 *****
BLACK	1192	15.9(0.1) [1%]	48.0(1.5) 191.7(1.6)	52.0(1.5) 191.3(1.6)	0.0 *****
HISPANIC	1182	9.5(0.2) [2%]	52.6(1.8) 200.5(1.6)	47.4(1.8) 201.2(1.7)	0.0 *****
ASIAN/PACIFIC AMERICAN	234	2.5(0.2) [9%]	50.1(3.7) 231.7(3.0)	49.9(3.7) 230.9(3.4)	0.0 *****
AMER IND/ALASKAN NATV	157	1.5(0.2) [12%]	58.8(5.6) 208.8(3.7)	41.2(5.6) 209.5(4.7)	0.0 *****
PARENTS' EDUCATION LEVEL					
LESS THAN H.S.	298	4.1(0.3) [7%]	45.6(3.6) 202.1(4.1)	54.4(3.6) 205.4(2.7)	0.0 *****
GRADUATED H.S.	852	12.3(0.5) [4%]	51.2(1.8) 211.7(1.9)	48.8(1.8) 214.6(2.4)	0.0 *****
SOME EDUC AFTER H.S.	522	7.3(0.4) [6%]	53.8(2.5) 225.9(1.9)	46.2(2.5) 220.7(2.3)	0.0 *****
GRADUATED COLLEGE	2880	40.9(1.0) [3%]	51.8(1.1) 225.9(1.2)	48.2(1.1) 225.3(1.4)	0.0 *****
UNKNOWN	2616	35.2(0.7) [2%]	47.9(1.2) 215.4(1.2)	52.1(1.2) 210.4(1.0)	0.0 *****

BEST COPY AVAILABLE

Table 16-113

Weighted Response Percentages and Proficiency Means (with Standard Errors) for Mathematics Main Samples, by Derived Race/ethnicity, Grade 4

	N	WEIGHTED PCT [CV]	WHITE	BLACK	HISPANIC	ASIAN	AMER IND	UNCLASSI	MISSING
-- TOTAL --	7176	100.0(0.0) [0%]	70.4(0.2) 226.8(0.9)	15.9(0.1) 191.5(1.3)	9.5(0.2) 200.8(1.4)	2.5(0.2) 231.3(2.4)	1.5(0.2) 209.1(3.2)	0.3(0.1) 207.7(4.2)	0.0 *****
GENDER									
MALE	3553	50.2(0.6) [1%]	70.3(0.5) 228.4(1.0)	15.2(0.4) 191.7(1.6)	9.9(0.3) 200.5(1.6)	2.5(0.3) 231.7(3.0)	1.8(0.2) 208.8(3.7)	0.4(0.1) 206.8(5.5)	0.0 *****
FEMALE	3623	49.8(0.6) [1%]	70.5(0.5) 225.1(1.3)	16.6(0.5) 191.3(1.6)	9.0(0.4) 201.2(1.7)	2.5(0.3) 230.9(3.4)	1.2(0.2) 209.5(4.7)	0.2(0.1) 209.1(6.5)	0.0 *****
RACE/ETHNICITY									
WHITE	4381	70.4(0.2) [0%]	100.0(0.0) 226.8(0.9)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
BLACK	1192	15.9(0.1) [1%]	0.0(0.0) *****(0.0)	100.0(0.0) 191.5(1.3)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
HISPANIC	1182	9.5(0.2) [2%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 200.8(1.4)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
ASIAN/PACIFIC AMERICAN	234	2.5(0.2) [9%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 231.3(2.4)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
AMER IND/ALASKAN NATV	157	1.5(0.2) [12%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 209.1(3.2)	0.0(0.0) *****(0.0)	0.0 *****
PARENTS' EDUCATION LEVEL									
LESS THAN H.S.	298	4.1(0.3) [7%]	67.1(3.0) 210.7(3.3)	13.6(2.5) 184.5(5.4)	16.8(2.0) 192.0(3.8)	0.5(0.4) 223.8(10.5)	2.0(0.9) 204.5(5.4)	0.0(0.0) *****(0.0)	0.0 *****
GRADUATED H.S.	852	12.3(0.5) [4%]	67.4(1.9) 221.1(1.8)	21.8(1.6) 190.6(3.0)	8.4(0.8) 207.5(3.5)	0.8(0.2) 227.4(5.7)	1.4(0.5) 210.9(7.4)	0.3(0.1) 202.5(5.5)	0.0 *****
SOME EDUC AFTER H.S.	522	7.3(0.4) [6%]	71.5(2.1) 232.6(1.9)	17.1(1.9) 193.9(3.6)	7.1(0.9) 206.4(4.0)	1.8(0.5) 230.3(9.5)	2.0(0.6) 207.0(7.2)	0.5(0.3) 223.1(9.8)	0.0 *****
GRADUATED COLLEGE	2880	40.9(1.0) [3%]	73.6(0.9) 234.1(1.1)	15.4(0.7) 192.9(1.9)	7.0(0.5) 206.7(2.4)	2.4(0.3) 238.7(3.1)	1.2(0.2) 210.4(5.1)	0.4(0.1) 212.9(6.6)	0.0 *****
UNKNOWN	2616	35.2(0.7) [2%]	67.9(0.9) 220.1(1.1)	14.4(0.7) 190.4(1.9)	12.4(0.5) 196.2(1.6)	3.4(0.4) 225.8(3.3)	1.7(0.3) 208.7(4.4)	0.2(0.1) 186.7(10.7)	0.0 *****

Table 16-114

Weighted Response Percentages and Proficiency Means (with Standard Errors) for Mathematics Main Samples, by Parents' Education Level, Grade 4

	N	WEIGHTED PCT [CV]	NOT HS	GRAD HS	POST HS	GRAD COL	UNKNOWN	MISSING
-- TOTAL --	7168	100.0(0.0) [0%]	4.1(0.3)	12.4(0.5)	7.3(0.4)	40.9(1.0)	35.3(0.7)	0.1
			203.9(2.6)	213.1(1.5)	223.5(1.5)	225.6(1.0)	212.8(0.8)	200.8
GENDER								
MALE	3551	50.2(0.6) [1%]	3.7(0.4)	12.6(0.8)	7.9(0.6)	42.2(1.1)	33.6(0.9)	0.0
			202.1(4.1)	211.7(1.9)	225.9(1.9)	225.9(1.2)	215.4(1.2)	219.1
FEMALE	3617	49.8(0.6) [1%]	4.5(0.4)	12.1(0.6)	6.8(0.5)	39.6(1.4)	36.9(1.1)	0.2
			205.4(2.7)	214.6(2.4)	220.7(2.3)	225.3(1.4)	210.4(1.0)	195.0
RACE/ETHNICITY								
WHITE	4376	70.4(0.2) [0%]	3.9(0.4)	11.8(0.7)	7.5(0.5)	42.8(1.3)	34.0(1.0)	0.1
			210.7(3.3)	221.1(1.8)	232.6(1.9)	234.1(1.1)	220.1(1.1)	199.2
BLACK	1191	15.9(0.1) [1%]	3.5(0.6)	16.9(1.4)	7.9(1.0)	39.7(2.0)	31.9(1.6)	0.1
			184.5(5.4)	190.6(3.0)	193.9(3.6)	192.9(1.9)	190.4(1.9)	202.8
HISPANIC	1181	9.5(0.2) [2%]	7.3(0.8)	10.9(1.1)	5.5(0.7)	30.1(1.7)	46.2(1.9)	0.1
			192.0(3.8)	207.5(3.5)	206.4(4.0)	206.7(2.4)	196.2(1.6)	187.7
ASIAN/PACIFIC AMERICAN	233	2.5(0.2) [9%]	0.9(0.6)	4.0(1.1)	5.5(1.7)	40.7(4.1)	49.0(4.0)	0.3
			223.8(10.5)	227.4(5.7)	230.3(9.5)	238.7(3.1)	225.8(3.3)	230.0
AMER IND/ALASKAN NATV	157	1.5(0.2) [12%]	5.4(2.2)	11.8(3.8)	9.6(2.5)	32.1(3.9)	41.1(4.9)	0.0
			204.5(5.4)	210.9(7.4)	207.0(7.2)	210.4(5.1)	208.7(4.4)	*****
PARENTS' EDUCATION LEVEL								
LESS THAN H.S.	298	4.1(0.3) [7%]	100.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0
			203.9(2.6)	*****(0.0)	*****(0.0)	*****(0.0)	*****(0.0)	*****
GRADUATED H.S.	852	12.4(0.5) [4%]	0.0(0.0)	100.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0
			*****(0.0)	213.1(1.5)	*****(0.0)	*****(0.0)	*****(0.0)	*****
SOME EDUC AFTER H.S.	522	7.3(0.4) [6%]	0.0(0.0)	0.0(0.0)	100.0(0.0)	0.0(0.0)	0.0(0.0)	0.0
			*****(0.0)	*****(0.0)	223.5(1.5)	*****(0.0)	*****(0.0)	*****
GRADUATED COLLEGE	2880	40.9(1.0) [3%]	0.0(0.0)	0.0(0.0)	0.0(0.0)	100.0(0.0)	0.0(0.0)	0.0
			*****(0.0)	*****(0.0)	*****(0.0)	225.6(1.0)	*****(0.0)	*****
UNKNOWN	2616	35.3(0.7) [2%]	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	100.0(0.0)	0.0
			*****(0.0)	*****(0.0)	*****(0.0)	*****(0.0)	212.8(0.8)	*****

Table 16-115

Weighted Response Percentages and Proficiency Means (with Standard Errors) for Mathematics Main Samples, by Gender, Grade 8

	N	WEIGHTED PCT [CV]	MALE	FEMALE	MISSING
-- TOTAL --	7663	100.0(0.0) [0%]	51.3(0.6) 267.4(1.1)	48.7(0.6) 268.0(1.0)	0.0 *****
GENDER					
MALE	3820	51.3(0.6) [1%]	100.0(0.0) 267.4(1.1)	0.0(0.0) *****(0.0)	0.0 *****
FEMALE	3843	48.7(0.6) [1%]	0.0(0.0) *****(0.0)	100.0(0.0) 268.0(1.0)	0.0 *****
RACE/ETHNICITY					
WHITE	4903	69.9(0.2) [0%]	51.0(0.7) 277.1(1.2)	49.0(0.7) 277.3(1.1)	0.0 *****
BLACK	1195	15.5(0.1) [1%]	51.5(1.7) 236.7(1.9)	48.5(1.7) 236.9(1.5)	0.0 *****
HISPANIC	1120	10.1(0.2) [2%]	51.5(1.8) 245.5(1.7)	48.5(1.8) 247.2(1.9)	0.0 *****
ASIAN/PACIFIC AMERICAN	281	2.5(0.2) [9%]	54.0(3.2) 287.3(6.4)	46.0(3.2) 288.7(5.6)	0.0 *****
AMER IND/ALASKAN NATV	112	1.3(0.2) [16%]	62.4(5.0) 251.9(3.9)	37.6(5.0) 258.3(4.1)	0.0 *****
PARENTS' EDUCATION LEVEL					
LESS THAN H.S.	563	7.6(0.5) [7%]	38.2(1.9) 253.4(2.3)	61.8(1.9) 245.5(2.0)	0.0 *****
GRADUATED H.S.	1679	23.6(0.7) [3%]	52.3(1.3) 255.5(1.7)	47.7(1.3) 257.7(1.6)	0.0 *****
SOME EDUC AFTER H.S.	1382	18.1(0.5) [3%]	48.1(1.7) 270.1(1.5)	51.9(1.7) 270.7(1.9)	0.0 *****
GRADUATED COLLEGE	3290	41.7(1.3) [3%]	53.5(1.2) 277.9(1.7)	46.5(1.2) 282.5(1.3)	0.0 *****
UNKNOWN	726	8.7(0.4) [5%]	56.5(2.2) 254.2(2.8)	43.5(2.2) 246.8(1.8)	0.0 *****

Table 16-116

Weighted Response Percentages and Proficiency Means (with Standard Errors)
for Mathematics Main Samples, by Derived Race/ethnicity, Grade 8

	N	WEIGHTED PCT [CV]	WHITE	BLACK	HISPANIC	ASIAN	AMER IND	UNCLASSI	MISSING
-- TOTAL --	7663	100.0(0.0) [0%]	69.9(0.2) 277.2(1.0)	15.5(0.1) 236.8(1.4)	10.1(0.2) 246.3(1.2)	2.5(0.2) 288.0(5.5)	1.3(0.2) 254.3(2.8)	0.6(0.2) 261.3(5.2)	0.0 *****
GENDER									
MALE	3820	51.3(0.6) [1%]	69.5(0.7) 277.1(1.2)	15.6(0.5) 236.7(1.9)	10.1(0.4) 245.5(1.7)	2.7(0.3) 287.3(6.4)	1.6(0.3) 251.9(3.9)	0.5(0.2) 257.2(5.4)	0.0 *****
FEMALE	3843	48.7(0.6) [1%]	70.3(0.6) 277.3(1.1)	15.5(0.5) 236.9(1.5)	10.1(0.4) 247.2(1.9)	2.4(0.3) 288.7(5.6)	1.0(0.2) 258.3(4.1)	0.7(0.3) 264.2(6.6)	0.0 *****
RACE/ETHNICITY									
WHITE	4903	69.9(0.2) [0%]	100.0(0.0) 277.2(1.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
BLACK	1195	15.5(0.1) [1%]	0.0(0.0) *****(0.0)	100.0(0.0) 236.8(1.4)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
HISPANIC	1120	10.1(0.2) [2%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 246.3(1.2)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
ASIAN/PACIFIC AMERICAN	281	2.5(0.2) [9%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 288.0(5.5)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
AMER IND/ALASKAN NATV	112	1.3(0.2) [16%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 254.3(2.8)	0.0(0.0) *****(0.0)	0.0 *****
PARENTS' EDUCATION LEVEL									
LESS THAN H.S.	563	7.6(0.5) [7%]	57.7(3.3) 256.7(2.0)	13.9(2.3) 229.1(3.9)	23.3(2.1) 238.9(2.6)	1.2(0.5) 250.7(9.9)	1.8(0.7) 254.7(8.9)	2.1(1.1) 251.1(8.6)	0.0 *****
GRADUATED H.S.	1679	23.6(0.7) [3%]	68.3(1.2) 265.5(1.4)	18.3(1.1) 231.2(1.8)	11.2(0.6) 244.2(2.1)	0.7(0.2) 275.3(7.1)	1.5(0.4) 245.4(5.9)	0.1(0.1) 255.6(12.9)	0.0 *****
SOME EDUC AFTER H.S.	1382	18.1(0.5) [3%]	72.4(1.6) 278.7(1.3)	16.6(1.2) 241.5(2.5)	8.5(0.8) 255.0(3.3)	1.1(0.3) 289.1(10.3)	0.9(0.2) 263.6(5.7)	0.6(0.3) 261.9(8.0)	0.0 *****
GRADUATED COLLEGE	3290	41.7(1.3) [3%]	76.2(0.8) 287.3(1.3)	12.9(0.6) 242.6(2.2)	5.5(0.4) 258.6(3.0)	3.7(0.6) 299.0(5.0)	1.1(0.2) 259.6(5.3)	0.5(0.2) 271.3(7.9)	0.0 *****
UNKNOWN	726	8.7(0.4) [5%]	50.7(2.1) 263.3(2.1)	18.8(2.1) 230.9(2.1)	20.8(1.4) 234.5(1.9)	6.3(1.2) 265.8(4.7)	2.4(0.5) 249.7(7.9)	1.1(0.9) 255.5(8.8)	0.0 *****

Table 16-117

Weighted Response Percentages and Proficiency Means (with Standard Errors) for Mathematics Main Samples, by Parents' Education Level, Grade 8

	N	WEIGHTED PCT [CV]	NOT HS	GRAD HS	POST HS	GRAD COL	UNKNOWN	MISSING
-- TOTAL --	7640	100.0(0.0) [0%]	7.6(0.5) 248.5(1.7)	23.7(0.7) 256.6(1.2)	18.2(0.5) 270.4(1.1)	41.8(1.3) 280.0(1.2)	8.7(0.4) 251.0(1.6)	0.3 235.8
GENDER								
MALE	3807	51.3(0.6) [1%]	5.7(0.5) 253.4(2.3)	24.1(1.0) 255.5(1.7)	17.0(0.8) 270.1(1.5)	43.6(1.5) 277.9(1.7)	9.6(0.6) 254.2(2.8)	0.3 233.0
FEMALE	3833	48.7(0.6) [1%]	9.7(0.7) 245.5(2.0)	23.2(0.8) 257.7(1.6)	19.4(0.8) 270.7(1.9)	39.9(1.5) 282.5(1.3)	7.8(0.6) 246.8(1.8)	0.3 239.6
RACE/ETHNICITY								
WHITE	4894	70.0(0.2) [0%]	6.3(0.7) 256.7(2.0)	23.1(0.9) 265.5(1.4)	18.8(0.7) 278.7(1.3)	45.5(1.5) 287.3(1.3)	6.3(0.4) 263.3(2.1)	0.2 265.5
BLACK	1187	15.5(0.1) [1%]	6.8(1.1) 229.1(3.9)	28.1(1.7) 231.2(1.8)	19.4(1.5) 241.5(2.5)	35.0(1.7) 242.6(2.2)	10.6(1.3) 230.9(2.1)	0.9 213.8
HISPANIC	1114	10.1(0.2) [2%]	17.6(1.7) 238.9(2.6)	26.3(1.5) 244.2(2.1)	15.2(1.4) 255.0(3.3)	22.9(1.6) 258.6(3.0)	18.0(1.2) 234.5(1.9)	0.4 221.9
ASIAN/PACIFIC AMERICAN	281	2.6(0.2) [9%]	3.5(1.5) 250.7(9.9)	6.2(1.6) 275.3(7.1)	7.8(2.1) 289.1(10.3)	61.1(6.6) 299.0(5.0)	21.5(4.5) 265.8(4.7)	0.0 *****
AMER IND/ALASKAN NATV	112	1.3(0.2) [16%]	10.8(3.4) 254.7(8.9)	26.4(4.5) 245.4(5.9)	12.8(2.2) 263.6(5.7)	34.1(4.6) 259.6(5.3)	16.0(3.1) 249.7(7.9)	0.0 *****
PARENTS' EDUCATION LEVEL								
LESS THAN H.S.	563	7.6(0.5) [7%]	100.0(0.0) 248.5(1.7)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
GRADUATED H.S.	1679	23.7(0.7) [3%]	0.0(0.0) *****(0.0)	100.0(0.0) 256.6(1.2)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
SOME EDUC AFTER H.S.	1382	18.2(0.5) [3%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 270.4(1.1)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
GRADUATED COLLEGE	3290	41.8(1.3) [3%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 280.0(1.2)	0.0(0.0) *****(0.0)	0.0 *****
UNKNOWN	726	8.7(0.4) [5%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 251.0(1.6)	0.0 *****

Table 16-118

Weighted Response Percentages and Proficiency Means (with Standard Errors) for Mathematics Main Samples, by Gender, Grade 12

	N	WEIGHTED PCT [CV]	MALE	FEMALE	MISSING
-- TOTAL --	6973	100.0(0.0) [1%]	48.8(0.8)	51.2(0.8)	0.0
			300.6(1.1)	297.0(1.0)	*****
GENDER					
MALE	3290	48.8(0.8) [2%]	100.0(0.0)	0.0(0.0)	0.0
			300.6(1.1)	*****(0.0)	*****
FEMALE	3683	51.2(0.8) [2%]	0.0(0.0)	100.0(0.0)	0.0
			*****(0.0)	297.0(1.0)	*****
RACE/ETHNICITY					
WHITE	4738	71.1(0.6) [1%]	49.5(0.9)	50.5(0.9)	0.0
			307.0(1.0)	303.1(1.0)	*****
BLACK	1122	14.7(0.4) [3%]	45.6(2.2)	54.4(2.2)	0.0
			276.7(2.3)	273.2(1.8)	*****
HISPANIC	734	9.5(0.5) [6%]	49.9(2.5)	50.1(2.5)	0.0
			280.9(3.4)	284.9(2.5)	*****
ASIAN/PACIFIC AMERICAN	310	3.9(0.2) [6%]	46.0(3.6)	54.0(3.6)	0.0
			317.6(4.2)	313.3(4.6)	*****
AMER IND/ALASKAN NATV	49	0.6(0.1) [15%]	45.2(12.4)	54.8(12.4)	0.0
			295.1(8.8)	269.6(9.4)	*****
PARENTS' EDUCATION LEVEL					
LESS THAN H.S.	495	6.5(0.4) [7%]	44.7(3.2)	55.3(3.2)	0.0
			281.1(2.3)	274.5(2.5)	*****
GRADUATED H.S.	1429	21.4(0.8) [4%]	48.3(1.4)	51.7(1.4)	0.0
			288.7(2.0)	285.4(1.5)	*****
SOME EDUC AFTER H.S.	1793	26.2(0.7) [3%]	45.4(1.5)	54.6(1.5)	0.0
			299.8(1.5)	296.1(1.4)	*****
GRADUATED COLLEGE	3048	42.9(1.1) [2%]	51.1(1.2)	48.9(1.2)	0.0
			311.5(1.5)	308.7(1.3)	*****
UNKNOWN	181	2.6(0.3) [10%]	57.1(3.9)	42.9(3.9)	0.0
			273.0(3.1)	279.0(5.4)	*****

Table 16-119

Weight Response Percentages and Proficiency Means (with Standard Errors) for Mathematics Main Samples, by Derived Race/ethnicity, Grade 12

	N	WEIGHTED PCT [CV]	WHITE	BLACK	HISPANIC	ASIAN	AMER IND	UNCLASSI	MISSING
-- TOTAL --	6973	100.0(0.0) [1%]	71.1(0.6) 305.0(0.9)	14.7(0.4) 274.8(1.7)	9.5(0.5) 282.9(1.8)	3.9(0.2) 315.3(3.5)	0.6(0.1) 281.1(9.0)	0.2(0.1) 300.2(9.7)	0.0 *****
GENDER									
MALE	3290	48.8(0.8) [2%]	72.1(1.0) 307.0(1.0)	13.8(0.8) 276.7(2.3)	9.7(0.7) 280.9(3.4)	3.7(0.4) 317.6(4.2)	0.5(0.2) 295.1(8.8)	0.2(0.1) 296.6(11.0)	0.0 *****
FEMALE	3683	51.2(0.8) [2%]	70.2(0.9) 303.1(1.0)	15.7(0.6) 273.2(1.8)	9.2(0.7) 284.9(2.5)	4.1(0.4) 313.3(4.6)	0.6(0.1) 269.6(9.4)	0.2(0.1) 304.2(13.9)	0.0 *****
RACE/ETHNICITY									
WHITE	4738	71.1(0.6) [1%]	100.0(0.0) 305.0(0.9)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
BLACK	1122	14.7(0.4) [3%]	0.0(0.0) *****(0.0)	100.0(0.0) 274.8(1.7)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
HISPANIC	734	9.5(0.5) [6%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 282.9(1.8)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
ASIAN/PACIFIC AMERICAN	310	3.9(0.2) [6%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 315.3(3.5)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
AMER IND/ALASKAN NATV	49	0.6(0.1) [15%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 281.1(9.0)	0.0(0.0) *****(0.0)	0.0 *****
PARENTS' EDUCATION LEVEL									
LESS THAN H.S.	495	6.5(0.4) [7%]	46.4(3.4) 283.0(2.6)	12.9(1.8) 263.3(5.0)	35.8(3.5) 274.7(2.9)	3.0(1.0) 299.3(7.6)	1.0(0.5) 241.5(10.2)	0.9(0.5) 272.9(15.6)	0.0 *****
GRADUATED H.S.	1429	21.4(0.8) [4%]	70.4(1.8) 292.9(1.3)	18.3(1.5) 265.6(2.3)	8.0(0.9) 278.1(2.3)	2.7(0.7) 304.2(6.7)	0.6(0.3) 291.2(15.0)	0.0(0.0) 294.1(****)	0.0 *****
SOME EDUC AFTER H.S.	1793	26.2(0.7) [3%]	71.4(1.0) 303.0(1.1)	16.6(0.8) 277.0(2.4)	8.2(0.9) 290.0(3.4)	3.0(0.5) 310.4(3.9)	0.7(0.2) 288.8(11.2)	0.1(0.1) 308.0(16.4)	0.0 *****
GRADUATED COLLEGE	3048	42.9(1.1) [2%]	77.1(0.7) 314.9(1.2)	11.6(0.6) 283.0(2.7)	6.2(0.4) 289.6(3.8)	4.6(0.4) 325.8(3.6)	0.3(0.1) 299.1(7.6)	0.3(0.1) 319.8(14.2)	0.0 *****
UNKNOWN	181	2.6(0.3) [10%]	36.0(5.1) 276.0(6.9)	23.2(5.3) 270.1(5.8)	24.2(3.9) 274.0(4.5)	13.4(3.6) 293.5(5.8)	2.6(1.4) 246.9(19.2)	0.6(0.6) 248.2(****)	0.0 *****

Table 16-120

Weight Response Percentages and Proficiency Means (with Standard Errors) for Mathematics Main Samples, by Parents' Education Level, Grade 12

	N	WEIGHTED PCT [CV]	NOT HS	GRAD HS	POST HS	GRAD COL	UNKNOWN	MISSING
-- TOTAL --	6946	100.0(0.0) [1%]	6.5(0.4) 277.5(1.7)	21.5(0.8) 287.0(1.4)	26.3(0.7) 297.8(1.0)	43.1(1.1) 310.1(1.2)	2.7(0.3) 275.6(3.0)	0.4 264.1
GENDER								
MALE	3274	48.8(0.8) [2%]	5.9(0.6) 281.1(2.3)	21.3(1.0) 288.7(2.0)	24.5(0.9) 299.8(1.5)	45.2(1.3) 311.5(1.5)	3.1(0.3) 273.0(3.1)	0.5 261.1
FEMALE	3672	51.2(0.8) [2%]	7.0(0.6) 274.5(2.5)	21.7(1.0) 285.4(1.5)	28.0(0.9) 296.1(1.4)	41.1(1.3) 308.7(1.3)	2.2(0.3) 279.0(5.4)	0.3 268.5
RACE/ETHNICITY								
WHITE	4719	71.1(0.6) [1%]	4.2(0.4) 283.0(2.6)	21.3(0.9) 292.9(1.3)	26.4(0.8) 303.0(1.1)	46.7(1.2) 314.9(1.2)	1.3(0.2) 276.0(6.9)	0.4 267.5
BLACK	1115	14.7(0.4) [3%]	5.7(0.7) 263.3(5.0)	26.7(2.1) 265.6(2.3)	29.6(1.6) 277.0(2.4)	33.8(2.1) 283.0(2.7)	4.2(1.1) 270.1(5.8)	0.5 245.5
HISPANIC	734	9.5(0.5) [6%]	24.4(2.4) 274.7(2.9)	18.1(1.9) 278.1(2.3)	22.7(2.0) 290.0(3.4)	28.0(2.7) 289.6(3.8)	6.7(0.9) 274.0(4.5)	0.0 *****
ASIAN/PACIFIC AMERICAN	310	3.9(0.2) [6%]	4.9(1.5) 299.3(7.6)	14.9(3.6) 304.2(6.7)	20.2(3.3) 310.4(3.9)	51.0(5.3) 325.8(3.6)	9.0(2.4) 293.5(5.8)	0.0 *****
AMER IND/ALASKAN NATV	48	0.6(0.1) [16%]	11.6(5.1) 241.5(10.2)	21.7(8.2) 291.2(15.0)	31.2(7.2) 288.8(11.2)	23.2(9.4) 299.1(7.6)	12.2(5.8) 246.9(19.2)	1.8 284.1
PARENTS' EDUCATION LEVEL								
LESS THAN H.S.	495	6.5(0.4) [7%]	100.0(0.0) 277.5(1.7)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
GRADUATED H.S.	1429	21.5(0.8) [4%]	0.0(0.0) *****(0.0)	100.0(0.0) 287.0(1.4)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
SOME EDUC AFTER H.S.	1793	26.3(0.7) [3%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 297.8(1.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0 *****
GRADUATED COLLEGE	3048	43.1(1.1) [2%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 310.1(1.2)	0.0(0.0) *****(0.0)	0.0 *****
UNKNOWN	181	2.7(0.3) [10%]	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	0.0(0.0) *****(0.0)	100.0(0.0) 275.6(3.0)	0.0 *****

Table 16-121
Proficiencies for Mathematics Long-term Trend Samples, Grade 4

	1978	1982	1986	1990	1992
-- TOTAL --	100.0(0.0) 218.6(0.8)*	100.0(0.0) 219.0(1.1)*	100.0(0.0) 221.7(1.0)*	100.0(0.0) 229.6(0.8)■	100.0(0.0) 229.6(0.8)•
GENDER					
MALE	49.7(0.5) 217.4(0.7)*	49.3(0.6) 217.1(1.2)*	50.0(0.6) 221.7(1.1)■*	49.4(0.6) 229.1(0.9)■	48.6(0.6) 230.8(1.0)•
FEMALE	50.3(0.5) 219.9(1.0)*	50.7(0.6) 220.8(1.2)*	50.0(0.6) 221.7(1.2)*	50.6(0.6) 230.2(1.1)■	51.4(0.6) 228.4(1.0)•
RACE/ETHNICITY					
WHITE	79.4(1.4) 224.1(0.9)*	78.6(2.5) 224.0(1.1)*	76.5(1.1) 226.9(1.1)*	74.5(1.1)■ 235.2(0.8)■	75.4(0.9)• 235.1(0.8)•
BLACK	13.8(1.4) 192.4(1.1)*	14.3(2.0) 194.9(1.6)*	14.9(0.5) 201.6(1.6)■	16.3(0.7) 208.4(2.2)■	15.9(0.4) 208.0(2.0)•
HISPANIC	5.4(0.7) 202.9(2.2)*	5.4(1.1) 204.0(1.3)*	6.2(1.1) 205.4(2.1)	5.5(0.6) 213.8(2.1)■	5.6(0.8) 211.9(2.3)•
OTHER	1.4(0.2)* 227.2(3.4)*	1.7(0.6) 238.5(3.4)	2.5(0.5) 221.8(7.5)	3.8(1.0) 235.2(3.2)	3.0(0.4)• 239.3(3.4)•
PARENTS' EDUCATION LEVEL					
LESS THAN H.S.	7.9(0.4)* 200.3(1.5)*	8.0(0.7)* 199.0(1.7)*	4.2(0.4)■ 200.6(2.5)*	4.9(0.4)■ 210.4(2.3)■	4.2(0.3)• 216.7(2.2)•
GRADUATED H.S.	23.0(0.8)* 219.2(1.1)	25.1(0.8)* 218.3(1.1)	16.4(0.7)■* 218.4(1.6)	16.0(0.7)■* 226.2(1.2)■	13.5(0.7)• 222.0(1.5)
SOME EDUC AFTER H.S.	8.8(0.4) 230.1(1.7)*	9.4(0.4)* 225.2(2.1)*	6.6(0.6)■ 228.6(2.1)*	7.4(0.4) 235.8(2.0)	7.8(0.4) 237.4(1.9)•
GRADUATED COLLEGE	23.6(1.1)* 231.3(1.1)*	30.1(1.5)■* 228.8(1.5)*	37.8(1.1)■ 231.3(1.1)*	40.1(1.1)■ 237.6(1.3)■	41.5(1.2)• 236.2(1.0)•
UNKNOWN	36.8(1.5) 211.4(1.1)*	27.3(1.1)■* 212.6(1.5)*	34.9(1.0) 214.3(1.4)*	31.7(0.8)■ 223.0(1.0)■	33.0(0.8)• 224.5(1.0)•

* Indicates a significant difference between 1992 and a previous year, at the 95% certainty level.

■ Indicates a significant difference between the earliest year with available data and a subsequent year, at the 95% certainty level. These notations (*,■) indicate statistical significance from a multiple comparison procedure based on the number of assessments with valid data.

• Indicates a significant difference between 1992 and the earliest year with availiable data, at the 95% certainty level. This notation indicates statistical significance without application of a multiple comparison procedure.

Table 16-122
Proficiencies for Mathematics Long-term Trend Samples, Grade 8

	1978	1982	1986	1990	1992
-- TOTAL --	100.0(0.0) 264.1(1.1)*	100.0(0.0) 268.6(1.1)■*	100.0(0.0) 269.0(1.2)■*	100.0(0.0) 270.4(0.9)■	100.0(0.0) 273.1(0.9)•
GENDER					
MALE	49.9(0.5) 263.6(1.3)*	50.2(0.4) 269.2(1.4)■*	49.8(0.9) 270.0(1.1)■*	49.8(0.6) 271.2(1.2)■	49.9(0.8) 274.1(1.1)•
FEMALE	50.1(0.5) 264.7(1.1)*	49.8(0.4) 268.0(1.1)*	50.2(0.9) 267.9(1.5)	50.2(0.6) 269.6(0.9)■	50.1(0.8) 272.0(1.0)•
RACE/ETHNICITY					
WHITE	80.2(1.7)* 271.6(0.8)*	79.2(2.1) 274.4(1.0)*	76.8(1.0) 273.6(1.3)*	73.4(0.7)■ 276.3(1.1)■	74.2(0.5)• 278.9(0.9)•
BLACK	13.1(1.5) 229.6(1.9)*	13.8(1.8) 240.4(1.6)■*	14.4(0.9) 249.2(2.3)■	15.6(0.3) 249.1(2.3)■	15.9(0.3) 250.2(1.9)•
HISPANIC	5.8(0.9) 238.0(2.0)*	5.0(1.2) 252.4(1.7)■*	6.6(1.1) 254.3(2.9)■	7.3(0.5) 254.6(1.8)■	7.0(0.5) 259.3(1.8)•
OTHER	0.9(0.2)* 272.5(3.5)	2.0(0.5) 274.5(4.1)	2.2(0.3)■ 282.7(3.4)	3.7(0.8)■ 273.5(7.2)	2.9(0.3)• 282.0(2.3)•
PARENTS' EDUCATION LEVEL					
LESS THAN H.S.	12.2(0.6)* 244.7(1.2)*	10.7(0.6)* 251.0(1.4)■*	7.9(1.1)■ 252.3(2.3)■	7.6(0.5)■ 253.4(1.8)■	5.9(0.5)• 255.5(1.0)•
GRADUATED H.S.	33.0(0.8)* 263.1(1.0)	34.4(0.8)* 262.9(0.8)	31.0(1.3)* 262.7(1.2)	26.8(0.8)■* 262.6(1.2)	23.1(0.9)• 263.2(1.2)
SOME EDUC AFTER H.S.	14.3(0.4)* 273.1(1.2)*	14.1(0.4)* 275.1(0.9)	15.6(0.6)* 273.7(0.8)*	16.8(0.6)■ 277.1(1.0)■	18.4(0.7)• 277.6(1.0)•
GRADUATED COLLEGE	25.7(1.2)* 283.8(1.2)	32.1(1.3)■* 282.3(1.5)	37.5(2.0)■* 279.9(1.4)	40.8(1.2)■ 280.4(1.0)	44.1(1.3)• 282.8(1.0)
UNKNOWN	14.8(0.9)* 239.5(1.3)*	8.8(0.8)■ 251.9(3.2)■	8.0(0.4)■ 247.4(2.3)■	7.9(0.5)■ 247.8(2.1)■	8.4(0.4)• 252.9(1.8)•

* Indicates a significant difference between 1992 and a previous year, at the 95% certainty level.

■ Indicates a significant difference between the earliest year with available data and a subsequent year, at the 95% certainty level. These notations (*,■) indicate statistical significance from a multiple comparison procedure based on the number of assessments with valid data.

• Indicates a significant difference between 1992 and the earliest year with availiable data, at the 95% certainty level. This notation indicates statistical significance without application of a multiple comparison procedure.

Table 16-123
Proficiencies for Mathematics Long-term Trend Samples, Grade 12

	1978	1982	1986	1990	1992
-- TOTAL --	100.0(0.0) 300.4(1.0)*	100.0(0.0) 298.5(0.9)*	100.0(0.0) 302.0(0.9)*	100.0(0.0) 304.6(0.9)▪	100.0(0.0) 306.7(0.9)•
GENDER					
MALE	48.7(0.5) 303.8(1.0)*	48.7(0.6) 301.5(1.0)*	49.0(1.2) 304.7(1.2)*	48.6(0.9) 306.3(1.1)	50.7(1.2) 308.9(1.1)•
FEMALE	51.3(0.5) 297.1(1.0)*	51.3(0.6) 295.6(1.0)*	51.0(1.2) 299.4(1.0)*	51.4(0.9) 302.9(1.1)▪	49.3(1.2) 304.5(1.1)•
RACE/ETHNICITY					
WHITE	83.1(1.3)* 305.9(0.9)*	80.7(2.0)* 303.7(0.9)*	77.6(0.5)▪* 307.5(1.0)*	73.3(0.5)▪ 309.5(1.0)▪	74.7(0.5)• 311.9(0.8)•
BLACK	11.8(1.1)* 268.4(1.3)*	12.5(1.7) 271.8(1.2)*	14.3(0.3) 278.6(2.1)▪	15.6(0.3)▪ 288.5(2.8)▪	14.8(0.3)• 285.8(2.2)•
HISPANIC	4.0(0.5)* 276.3(2.3)*	4.9(1.0) 276.7(1.8)*	5.5(0.3)* 283.1(2.9)	6.9(0.4)▪ 283.5(2.9)	7.4(0.5)• 292.2(2.6)•
OTHER	1.1(0.1)* 312.9(3.3)	1.9(0.6) 309.4(4.5)	2.6(0.4)▪ 304.7(7.2)	4.2(0.5)▪ 312.5(5.2)	3.1(0.2)• 317.1(4.0)
PARENTS' EDUCATION LEVEL					
LESS THAN H.S.	13.3(0.6)* 279.6(1.2)	13.8(0.9)* 279.3(1.0)	8.3(0.4)▪ 279.3(2.3)	7.9(0.6)▪ 285.4(2.2)	8.1(0.6)• 285.5(2.3)•
GRADUATED H.S.	33.3(0.7)* 293.9(0.8)	32.7(0.8)* 293.4(0.8)	27.9(1.1)▪* 293.1(1.0)	26.4(1.1)▪* 293.7(0.9)	21.4(0.9)• 297.6(1.7)
SOME EDUC AFTER H.S.	16.2(0.4)* 305.3(0.9)	17.7(0.5)* 303.9(0.9)*	24.1(1.0)▪ 305.2(1.2)	23.8(0.9)▪ 307.7(1.0)	25.4(0.9)• 307.5(1.1)
GRADUATED COLLEGE	32.4(1.1)* 316.8(1.0)	31.8(1.3)* 312.4(1.0)▪	36.9(1.2)▪* 313.9(1.4)	38.9(1.4)▪ 316.2(1.3)	42.6(1.4)• 315.9(1.0)
UNKNOWN	4.8(0.4)* 275.7(1.9)*	4.0(0.3)* 271.7(1.8)*	2.8(0.3)▪ 280.6(2.4)	3.0(0.4)▪ 276.8(2.8)*	2.5(0.3)• 290.2(3.9)•

* Indicates a significant difference between 1992 and a previous year, at the 95% certainty level.

▪ Indicates a significant difference between the earliest year with available data and a subsequent year, at the 95% certainty level. These notations (*,▪) indicate statistical significance from a multiple comparison procedure based on the number of assessments with valid data.

• Indicates a significant difference between 1992 and the earliest year with availiable data, at the 95% certainty level. This notation indicates statistical significance without application of a multiple comparison procedure.

Table 16-124
Proficiencies for Science Long-term Trend Samples, Grade 4

	1977	1982	1986	1990	1992
-- TOTAL --	100.0(0.0) 219.9(1.2)*	100.0(0.0) 220.8(1.8)*	100.0(0.0) 224.3(1.2)■*	100.0(0.0) 228.7(0.8)■	100.0(0.0) 230.6(1.0)•
GENDER					
MALE	50.6(0.4)* 222.1(1.3)*	50.5(1.5) 221.0(2.3)*	50.0(0.6) 227.3(1.4)■*	49.4(0.6) 230.3(1.1)■*	48.6(0.6)• 234.7(1.2)•
FEMALE	49.4(0.4)* 217.6(1.2)*	49.5(1.5) 220.7(2.0)*	50.0(0.6) 221.3(1.4)*	50.6(0.6) 227.1(1.0)■	51.4(0.6)• 226.7(1.0)•
RACE/ETHNICITY					
WHITE	79.6(1.6) 229.6(0.9)*	78.6(2.6) 229.0(1.9)*	76.5(1.1) 231.9(1.2)*	74.5(1.1)■ 237.5(0.8)■	75.4(0.9)• 239.1(1.0)•
BLACK	13.7(1.4) 174.8(1.8)*	14.3(2.1) 187.0(3.0)■*	14.9(0.5) 196.2(1.9)■	16.3(0.7) 196.4(2.0)■	15.9(0.4) 200.3(2.7)•
HISPANIC	5.3(0.9) 191.9(2.7)*	5.2(1.3) 189.0(4.2)*	6.2(1.1) 199.4(3.1)	5.5(0.6) 206.2(2.2)■	5.6(0.8) 204.7(2.8)•
OTHER	1.5(0.3)* 214.4(5.4)	2.0(1.0) 222.8(5.3)	2.5(0.5) 220.6(4.6)	3.8(1.0) 227.4(3.6)	3.0(0.4)• 226.5(3.4)
PARENTS' EDUCATION LEVEL					
LESS THAN H.S.	9.0(0.4)* 198.5(2.2)*	6.5(0.9) 198.2(6.0)*	4.2(0.4)■ 203.6(2.9)*	4.9(0.4)■ 209.8(2.7)■	4.2(0.3)• 217.2(2.6)•
GRADUATED H.S.	26.7(0.5)* 223.0(1.4)	14.7(1.1)■ 218.0(3.3)	16.4(0.7)■* 219.6(1.5)	16.0(0.7)■* 225.8(1.7)	13.5(0.7)• 222.0(1.9)
SOME EDUC AFTER H.S.	7.2(0.3) 237.2(1.5)	8.3(0.6) 229.1(3.2)	6.6(0.6) 235.8(2.6)	7.4(0.4) 237.6(2.1)	7.8(0.4) 236.6(2.4)
GRADUATED COLLEGE	23.1(0.7)* 232.3(1.4)*	42.0(2.3)■ 230.5(2.3)*	37.8(1.1)■ 235.2(1.4)	40.1(1.1)■ 236.2(1.3)	41.5(1.2)• 238.9(1.2)•
UNKNOWN	34.0(0.7) 211.0(1.4)*	28.5(1.8)■ 210.8(2.8)*	34.9(1.0) 215.3(1.5)*	31.7(0.8) 221.5(1.2)■	33.0(0.8) 224.2(1.4)•

* Indicates a significant difference between 1992 and a previous year, at the 95% certainty level.

■ Indicates a significant difference between the earliest year with available data and a subsequent year, at the 95% certainty level. These notations (*,■) indicate statistical significance from a multiple comparison procedure based on the number of assessments with valid data.

• Indicates a significant difference between 1992 and the earliest year with availiable data, at the 95% certainty level. This notation indicates statistical significance without application of a multiple comparison procedure.

BEST COPY AVAILABLE

Table 16-125
Proficiencies for Science Long-term Trend Samples, Grade 8

	1977	1982	1986	1990	1992
-- TOTAL --	100.0(0.0) 247.4(1.1)*	100.0(0.0) 250.1(1.3)*	100.0(0.0) 251.4(1.4)*	100.0(0.0) 255.2(0.9)■	100.0(0.0) 258.0(0.8)•
GENDER					
MALE	49.8(0.3) 251.1(1.3)*	48.5(0.7) 255.6(1.5)	49.8(0.9) 256.1(1.6)	49.8(0.6) 258.5(1.1)■	49.9(0.8) 260.1(1.2)•
FEMALE	50.2(0.3) 243.7(1.2)*	51.5(0.7) 245.0(1.3)*	50.2(0.9) 246.9(1.5)*	50.2(0.6) 251.8(1.1)■*	50.1(0.8) 256.0(1.0)•
RACE/ETHNICITY					
WHITE	80.4(1.6)* 256.1(0.8)*	79.2(2.1) 257.3(1.1)*	76.8(1.0) 259.2(1.4)*	73.4(0.7)■ 264.1(0.9)■	74.2(0.5)• 267.1(1.0)•
BLACK	13.0(1.2) 208.1(2.4)*	13.8(1.9) 217.1(1.3)■	14.4(0.9) 221.6(2.5)■	15.6(0.3) 225.7(3.1)■	15.9(0.3)• 224.4(2.7)•
HISPANIC	5.0(1.1) 213.4(1.9)*	5.3(1.0) 225.5(3.9)■*	6.6(1.1) 226.1(3.1)■*	7.3(0.5) 231.6(2.6)■	7.0(0.5) 237.5(2.6)•
OTHER	1.6(0.4)* 235.0(3.3)*	1.7(0.7) 262.4(5.9)■	2.2(0.3) 253.0(4.0)■	3.7(0.8) 248.2(10.9)	2.9(0.3)• 260.7(2.7)•
PARENTS' EDUCATION LEVEL					
LESS THAN H.S.	12.7(0.7)* 223.5(1.3)*	9.7(0.6)■* 225.3(1.9)	7.9(1.1)■ 229.4(2.7)	7.6(0.5)■ 232.9(2.1)■	5.9(0.5)• 233.8(2.9)•
GRADUATED H.S.	32.8(0.6)* 245.3(1.1)	25.6(1.1)■ 243.1(1.3)	31.0(1.3)* 244.8(1.4)	26.8(0.8)■* 247.3(1.3)	23.1(0.9)• 246.4(1.4)
SOME EDUC AFTER H.S.	15.0(0.5)* 260.3(1.3)*	16.8(0.6) 258.8(1.5)*	15.6(0.6)* 257.8(1.4)*	16.8(0.6) 262.8(1.2)	18.4(0.7)• 265.9(1.1)•
GRADUATED COLLEGE	26.6(1.0)* 266.4(1.0)	37.3(1.5)■* 263.5(1.5)*	37.5(2.0)■* 264.4(1.9)	40.8(1.2)■ 267.5(1.1)	44.1(1.3)• 269.2(1.0)•
UNKNOWN	12.9(1.1)* 221.9(1.8)*	10.5(1.2) 229.1(2.8)	8.0(0.4)■ 226.5(2.7)	7.9(0.5)■ 224.3(2.1)*	8.4(0.4)• 231.6(2.0)•

* Indicates a significant difference between 1992 and a previous year, at the 95% certainty level.

■ Indicates a significant difference between the earliest year with available data and a subsequent year, at the 95% certainty level. These notations (*,■) indicate statistical significance from a multiple comparison procedure based on the number of assessments with valid data.

• Indicates a significant difference between 1992 and the earliest year with availiable data, at the 95% certainty level. This notation indicates statistical significance without application of a multiple comparison procedure.

Table 16-126
Proficiencies for Science Long-term Trend Samples, Grade 12

	1977	1982	1986	1990	1992
-- TOTAL --	100.0(0.0) 289.5(1.0)*	100.0(0.0) 283.3(1.2)■*	100.0(0.0) 288.5(1.4)*	100.0(0.0) 290.4(1.1)	100.0(0.0) 294.1(1.3)•
GENDER					
MALE	49.5(0.7) 297.0(1.2)	48.4(0.7) 291.9(1.4)■*	49.0(1.2) 294.9(1.9)	48.6(0.9) 295.6(1.3)	50.7(1.2) 299.1(1.7)
FEMALE	50.5(0.7) 282.2(1.1)*	51.6(0.7) 275.2(1.3)■*	51.0(1.2) 282.3(1.5)*	51.4(0.9) 285.4(1.6)	49.3(1.2) 289.0(1.5)•
RACE/ETHNICITY					
WHITE	83.4(1.3)* 297.7(0.7)*	80.7(2.0)* 293.1(1.0)■*	77.6(0.5)■* 297.5(1.7)*	73.3(0.5)■ 300.9(1.1)■	74.6(0.5)• 304.2(1.3)•
BLACK	11.6(1.1)* 240.2(1.5)*	12.5(1.4) 234.7(1.7)*	14.3(0.3)■ 252.8(2.9)■	15.6(0.3)■ 253.0(4.5)■	14.8(0.3)• 256.2(3.2)•
HISPANIC	3.7(0.9)* 262.3(2.2)	4.5(1.1) 248.7(2.3)■*	5.5(0.3)* 259.3(3.8)	6.9(0.4)■ 261.5(4.4)	7.4(0.5)• 270.2(5.6)
OTHER	1.4(0.2)* 284.4(4.0)	2.2(0.7) 269.1(5.2)*	2.6(0.4)■ 276.8(11.2)	4.2(0.5)■ 292.0(5.6)	3.1(0.2)• 288.4(4.8)
PARENTS' EDUCATION LEVEL					
LESS THAN H.S.	15.2(0.9)* 265.3(1.3)	12.8(0.7)* 258.5(2.4)	8.3(0.4)■ 257.5(3.1)	7.9(0.6)■ 261.4(2.8)	8.1(0.6)• 262.0(3.8)
GRADUATED H.S.	33.1(0.6)* 284.4(0.8)	28.7(0.9)■* 275.2(1.6)■	27.9(1.1)■* 277.0(2.0)■	26.4(1.1)■* 276.3(1.4)■	21.4(0.9)• 280.2(2.4)
SOME EDUC AFTER H.S.	17.0(0.4)* 295.6(1.1)	21.5(0.6)■* 290.1(1.7)■	24.1(1.0)■ 295.1(2.5)	23.8(0.9)■ 296.5(1.6)	25.4(0.9)• 295.9(1.7)
GRADUATED COLLEGE	30.2(1.2)* 309.3(1.0)	32.4(1.4)* 300.2(1.7)■*	36.9(1.2)■* 303.8(2.1)	38.9(1.4)■ 305.5(1.7)	42.6(1.4)• 308.3(1.3)
UNKNOWN	4.4(0.4)* 252.6(3.2)	4.7(0.8)* 251.6(3.9)	2.8(0.3)■ 245.4(5.5)	3.0(0.4)■ 248.2(5.5)	2.5(0.3)• 257.6(7.4)

* Indicates a significant difference between 1992 and a previous year, at the 95% certainty level.

■ Indicates a significant difference between the earliest year with available data and a subsequent year, at the 95% certainty level. These notations (*,■) indicate statistical significance from a multiple comparison procedure based on the number of assessments with valid data.

• Indicates a significant difference between 1992 and the earliest year with availiable data, at the 95% certainty level. This notation indicates statistical significance without application of a multiple comparison procedure.

Table 16-127

Weight Response Percentages and Proficiency Means (with Standard Errors) for Writing Main Samples, by Gender, Grade 4

	N	WTD % (CV)	MALE	FEMALE	MISSING
--TOTAL--	7166	100.0 (0%)	49.3(0.5)	50.7(0.5)	0.0(0.0)
			215.4(1.2)	228.0(1.6)	0.0(0.0)
GENDER					
MALE	3483	49.3 (1%)	100.0(0.0)	0.0(0.0)	0.0(0.0)
			215.4(1.2)	0.0(0.0)	0.0(0.0)
FEMALE	3683	50.7 (1%)	0.0(0.0)	100.0(0.0)	0.0(0.0)
			0.0(0.0)	228.0(1.6)	0.0(0.0)
DERIVED ETHNICITY/RACE					
WHITE	4420	70.5 (0%)	49.6(0.7)	50.4(0.7)	0.0(0.0)
			222.4(1.5)	236.2(1.7)	0.0(0.0)
BLACK	1205	15.8 (1%)	45.7(1.5)	54.3(1.5)	0.0(0.0)
			189.4(2.9)	200.3(2.6)	0.0(0.0)
HISPANIC	1119	9.4 (2%)	52.3(1.9)	47.7(1.9)	0.0(0.0)
			202.3(2.4)	213.8(3.2)	0.0(0.0)
ASIAN AMERICAN	238	2.3 (10%)	51.0(3.3)	49.0(3.3)	0.0(0.0)
			226.5(4.2)	240.7(5.6)	0.0(0.0)
AMERICAN INDIAN	169	1.8 (12%)	46.7(4.8)	53.3(4.8)	0.0(0.0)
			204.6(4.8)	224.1(5.1)	0.0(0.0)
PARENTAL EDUCATION					
LESS THAN HS	305	4.0 (7%)	43.5(2.7)	56.5(2.7)	0.0(0.0)
			198.8(5.4)	215.2(4.5)	0.0(0.0)
GRADUATED HS	958	13.8 (4%)	49.9(1.7)	50.1(1.7)	0.0(0.0)
			207.5(2.2)	219.5(2.4)	0.0(0.0)
SOME EDU AFTER HS	619	9.2 (5%)	45.7(2.2)	54.3(2.2)	0.0(0.0)
			220.9(3.8)	233.9(3.1)	0.0(0.0)
GRADUATED COLLEGE	2856	40.5 (3%)	51.8(0.9)	48.2(0.9)	0.0(0.0)
			225.1(1.8)	238.6(1.6)	0.0(0.0)
UNKNOWN	2407	32.2 (2%)	47.3(1.1)	52.7(1.1)	0.0(0.0)
			206.4(1.6)	219.6(2.3)	0.0(0.0)

Table 16-128

Weight Response Percentages and Proficiency Means (with Standard Errors) for Writing Main Samples, by Derived Race/ethnicity, Grade 4

	N	WTD % (CV)	WHITE	BLACK	HISPANIC	ASIAN	AMER IND	UNCLASSI	MISSING
--TOTAL--	7166	100.0 (0%)	70.5(0.2)	15.8(0.1)	9.4(0.2)	2.3(0.2)	1.8(0.2)	0.2(0.1)	0.0(0.0)
			229.4(1.3)	195.3(2.3)	207.8(2.0)	233.4(3.9)	215.0(3.5)	233.0(7.6)	0.0(0.0)
GENDER									
MALE	3483	49.3 (1%)	71.0(0.6)	14.7(0.5)	10.0(0.4)	2.4(0.3)	1.7(0.3)	0.3(0.2)	0.0(0.0)
			222.4(1.5)	189.4(2.9)	202.3(2.4)	226.5(4.2)	204.6(4.8)	237.0(14.4)	0.0(0.0)
FEMALE	3683	50.7 (1%)	70.0(0.6)	16.9(0.4)	8.9(0.4)	2.2(0.3)	1.9(0.3)	0.1(0.1)	0.0(0.0)
			236.2(1.7)	200.3(2.6)	213.8(3.2)	240.7(5.6)	224.1(5.1)	223.6(18.0)	0.0(0.0)
DERIVED ETHNICITY/RACE									
WHITE	4420	70.5 (0%)	100.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
			229.4(1.3)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
BLACK	1205	15.8 (1%)	0.0(0.0)	100.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
			0.0(0.0)	195.3(2.3)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
HISPANIC	1119	9.4 (2%)	0.0(0.0)	0.0(0.0)	100.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
			0.0(0.0)	0.0(0.0)	207.8(2.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
ASIAN AMERICAN	238	2.3 (10%)	0.0(0.0)	0.0(0.0)	0.0(0.0)	100.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
			0.0(0.0)	0.0(0.0)	0.0(0.0)	233.4(3.9)	0.0(0.0)	0.0(0.0)	0.0(0.0)
AMERICAN INDIAN	169	1.8 (12%)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	100.0(0.0)	0.0(0.0)	0.0(0.0)
			0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	215.0(3.5)	0.0(0.0)	0.0(0.0)
PARENTAL EDUCATION									
LESS THAN HS	305	4.0 (7%)	58.1(3.5)	17.1(2.4)	20.5(2.4)	1.9(0.8)	2.3(0.7)	0.0(0.0)	0.0(0.0)
			213.5(4.9)	189.9(5.9)	207.2(6.3)	209.0(11.4)	212.4(16.8)	0.0(0.0)	0.0(0.0)
GRADUATED HS	958	13.8 (4%)	71.7(1.4)	17.2(1.2)	8.0(0.6)	0.9(0.3)	2.1(0.4)	0.1(0.1)	0.0(0.0)
			220.6(2.2)	190.7(2.9)	198.2(3.6)	227.7(10.0)	204.4(9.0)	271.6(****)	0.0(0.0)
SOME EDU AFTER HS	619	9.2 (5%)	77.7(1.7)	11.3(1.4)	7.0(0.9)	1.8(0.5)	2.2(0.6)	0.1(0.1)	0.0(0.0)
			233.1(2.9)	197.5(7.6)	217.5(4.7)	235.4(9.5)	233.0(6.9)	223.4(27.3)	0.0(0.0)
GRADUATED COLLEGE	2856	40.5 (3%)	73.2(0.8)	15.9(0.6)	6.5(0.4)	2.5(0.3)	1.7(0.3)	0.3(0.2)	0.0(0.0)
			238.7(1.2)	202.1(3.4)	219.9(3.5)	244.6(5.1)	224.4(5.0)	239.7(36.9)	0.0(0.0)
UNKNOWN	2407	32.2 (2%)	66.3(1.0)	16.2(0.8)	12.9(0.5)	2.9(0.3)	1.6(0.3)	0.2(0.1)	0.0(0.0)
			221.0(1.7)	189.6(3.4)	202.7(2.3)	223.7(5.0)	201.8(6.4)	217.7(20.3)	0.0(0.0)

Table 16-129

Weight Response Percentages and Proficiency Means (with Standard Errors) for Writing Main Samples, by Parents' Education Level, Grade 4

	N	WTD % (CV)	NOT HS	GRAD HS	POST HS	GRAD COL	UNKNOWN	MISSING
--TOTAL--	7145	100.0 (0%)	4.0(0.3)	13.9(0.5)	9.2(0.4)	40.6(1.0)	32.3(0.8)	0.2(0.1)
			208.0(3.5)	213.5(1.6)	228.0(2.5)	231.6(1.2)	213.4(1.6)	174.8(23.9)
GENDER								
MALE	3471	49.2 (1%)	3.6(0.3)	14.1(0.7)	8.5(0.5)	42.8(1.2)	31.0(1.0)	0.3(0.1)
			198.8(5.4)	207.5(2.2)	220.9(3.8)	225.1(1.8)	206.4(1.6)	171.0(19.8)
FEMALE	3674	50.8 (1%)	4.5(0.4)	13.7(0.7)	9.8(0.6)	38.5(1.1)	33.5(0.9)	0.2(0.1)
			215.2(4.5)	219.5(2.4)	233.9(3.1)	238.6(1.6)	219.6(2.3)	180.7(42.7)
DERIVED ETHNICITY/RACE								
WHITE	4413	70.6 (0%)	3.3(0.4)	14.1(0.6)	10.1(0.5)	42.1(1.3)	30.3(1.0)	0.2(0.1)
			213.5(4.9)	220.6(2.2)	233.1(2.9)	238.7(1.2)	221.0(1.7)	216.6(22.1)
BLACK	1202	15.8 (1%)	4.4(0.6)	15.1(1.2)	6.6(0.9)	40.8(1.7)	33.1(1.9)	0.2(0.1)
			189.9(5.9)	190.7(2.9)	197.5(7.6)	202.1(3.4)	189.6(3.4)	142.5(31.9)
HISPANIC	1109	9.4 (2%)	8.8(1.1)	11.8(0.9)	6.8(0.9)	28.3(1.9)	44.3(1.5)	0.7(0.4)
			207.2(6.3)	198.2(3.6)	217.5(4.7)	219.9(3.5)	202.7(2.3)	107.2(53.2)
ASIAN AMERICAN	238	2.3 (10%)	3.4(1.4)	5.5(1.7)	7.0(1.6)	44.1(3.3)	40.0(2.9)	0.0(0.0)
			209.0(11.4)	227.7(10.0)	235.4(9.5)	244.6(5.1)	223.7(5.0)	0.0(0.0)
AMERICAN INDIAN	169	1.8 (12%)	5.2(1.5)	16.2(2.9)	11.6(3.0)	38.1(3.6)	28.9(4.4)	0.0(0.0)
			212.4(16.8)	204.4(9.0)	233.0(6.9)	224.4(5.0)	201.8(6.4)	0.0(0.0)
PARENTAL EDUCATION								
LESS THAN HS	305	4.0 (7%)	100.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
			208.0(3.5)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
GRADUATED HS	958	13.9 (4%)	0.0(0.0)	100.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
			0.0(0.0)	213.5(1.6)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
SOME EDU AFTER HS	619	9.2 (5%)	0.0(0.0)	0.0(0.0)	100.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
			0.0(0.0)	0.0(0.0)	228.0(2.5)	0.0(0.0)	0.0(0.0)	0.0(0.0)
GRADUATED COLLEGE	2856	40.6 (3%)	0.0(0.0)	0.0(0.0)	0.0(0.0)	100.0(0.0)	0.0(0.0)	0.0(0.0)
			0.0(0.0)	0.0(0.0)	0.0(0.0)	231.6(1.2)	0.0(0.0)	0.0(0.0)
UNKNOWN	2407	32.3 (2%)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	100.0(0.0)	0.0(0.0)
			0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	213.4(1.6)	0.0(0.0)

Table 16-130

Weight Response Percentages and Proficiency Means (with Standard Errors)
for Writing Main Samples, by Gender, Grade 8

	N	WTD % (CV)	MALE	FEMALE	MISSING
--TOTAL--	11112	100.0 (0%)	50.4(0.5)	49.6(0.5)	0.0(0.0)
			251.5(1.1)	271.8(1.2)	0.0(0.0)
GENDER					
MALE	5583	50.4 (1%)	100.0(0.0)	0.0(0.0)	0.0(0.0)
			251.5(1.1)	0.0(0.0)	0.0(0.0)
FEMALE	5529	49.6 (1%)	0.0(0.0)	100.0(0.0)	0.0(0.0)
			0.0(0.0)	271.8(1.2)	0.0(0.0)
DERIVED ETHNICITY/RACE					
WHITE	7083	69.8 (0%)	50.0(0.6)	50.0(0.6)	0.0(0.0)
			257.8(1.1)	278.3(1.4)	0.0(0.0)
BLACK	1785	15.6 (1%)	49.6(1.5)	50.4(1.5)	0.0(0.0)
			232.6(2.6)	250.7(2.3)	0.0(0.0)
HISPANIC	1621	10.2 (2%)	52.9(1.5)	47.1(1.5)	0.0(0.0)
			238.9(2.5)	258.1(2.1)	0.0(0.0)
ASIAN AMERICAN	421	2.7 (7%)	55.9(2.4)	44.1(2.4)	0.0(0.0)
			258.0(5.2)	279.3(4.9)	0.0(0.0)
AMERICAN INDIAN	164	1.4 (13%)	53.0(4.5)	47.0(4.5)	0.0(0.0)
			232.8(6.2)	266.1(5.5)	0.0(0.0)
PARENTAL EDUCATION					
LESS THAN HS	813	7.6 (5%)	41.1(1.5)	58.9(1.5)	0.0(0.0)
			237.4(2.3)	259.6(2.7)	0.0(0.0)
GRADUATED HS	2534	24.5 (3%)	49.9(1.1)	50.1(1.1)	0.0(0.0)
			243.3(2.3)	265.5(2.2)	0.0(0.0)
SOME EDU AFTER HS	2133	19.6 (3%)	48.5(1.2)	51.5(1.2)	0.0(0.0)
			254.1(1.3)	274.4(1.5)	0.0(0.0)
GRADUATED COLLEGE	4670	40.2 (3%)	52.3(0.8)	47.7(0.8)	0.0(0.0)
			261.5(1.5)	281.0(1.5)	0.0(0.0)
UNKNOWN	936	7.9 (5%)	56.0(1.6)	44.0(1.6)	0.0(0.0)
			231.6(2.8)	252.5(2.8)	0.0(0.0)

Table 16-131

Weight Response Percentages and Proficiency Means (with Standard Errors) for Writing Main Samples, by Derived Race/ethnicity, Grade 8

	N	WTD % (CV)	WHITE	BLACK	HISPANIC	ASIAN	AMER IND	UNCLASSI	MISSING
--TOTAL--	11112	100.0 (0%)	69.8(0.2)	15.6(0.1)	10.2(0.2)	2.7(0.2)	1.4(0.2)	0.3(0.1)	0.0(0.0)
			268.0(1.1)	241.7(2.2)	247.9(1.7)	267.4(4.6)	248.5(4.3)	265.6(9.5)	0.0(0.0)
GENDER									
MALE	5583	50.4 (1%)	69.3(0.6)	15.3(0.4)	10.7(0.4)	2.9(0.3)	1.5(0.2)	0.3(0.1)	0.0(0.0)
			257.8(1.1)	232.6(2.6)	238.9(2.5)	258.0(5.2)	232.8(6.2)	256.8(10.1)	0.0(0.0)
FEMALE	5529	49.6 (1%)	70.4(0.5)	15.8(0.4)	9.7(0.3)	2.4(0.2)	1.4(0.2)	0.4(0.1)	0.0(0.0)
			278.3(1.4)	250.7(2.3)	258.1(2.1)	279.3(4.9)	266.1(5.5)	272.3(10.7)	0.0(0.0)
DERIVED ETHNICITY/RACE									
WHITE	7083	69.8 (0%)	100.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
			268.0(1.1)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
BLACK	1785	15.6 (1%)	0.0(0.0)	100.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
			0.0(0.0)	241.7(2.2)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
HISPANIC	1621	10.2 (2%)	0.0(0.0)	0.0(0.0)	100.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
			0.0(0.0)	0.0(0.0)	247.9(1.7)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
ASIAN AMERICAN	421	2.7 (7%)	0.0(0.0)	0.0(0.0)	0.0(0.0)	100.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
			0.0(0.0)	0.0(0.0)	0.0(0.0)	267.4(4.6)	0.0(0.0)	0.0(0.0)	0.0(0.0)
AMERICAN INDIAN	164	1.4 (13%)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	100.0(0.0)	0.0(0.0)	0.0(0.0)
			0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	248.5(4.3)	0.0(0.0)	0.0(0.0)
PARENTAL EDUCATION									
LESS THAN HS	813	7.6 (5%)	56.1(2.2)	14.5(1.3)	25.3(1.6)	1.3(0.3)	1.8(0.6)	1.0(0.4)	0.0(0.0)
			255.6(2.4)	240.8(4.6)	245.6(3.1)	262.4(14.1)	229.1(16.1)	251.4(12.6)	0.0(0.0)
GRADUATED HS	2534	24.5 (3%)	69.3(1.1)	17.9(0.8)	10.4(0.6)	0.9(0.2)	1.4(0.3)	0.2(0.1)	0.0(0.0)
			260.7(2.2)	235.8(3.9)	245.1(2.1)	263.3(8.2)	244.4(5.5)	248.6(28.1)	0.0(0.0)
SOME EDU AFTER HS	2133	19.6 (3%)	73.0(1.1)	16.2(0.9)	7.4(0.5)	1.6(0.2)	1.6(0.3)	0.2(0.1)	0.0(0.0)
			269.3(1.3)	248.7(2.7)	254.5(4.7)	259.8(8.4)	256.5(5.7)	309.8(19.4)	0.0(0.0)
GRADUATED COLLEGE	4670	40.2 (3%)	75.4(0.7)	13.0(0.5)	6.2(0.4)	3.8(0.4)	1.2(0.2)	0.3(0.1)	0.0(0.0)
			276.2(1.4)	245.8(2.9)	257.6(3.2)	276.6(5.2)	254.9(7.0)	274.6(9.4)	0.0(0.0)
UNKNOWN	936	7.9 (5%)	48.8(2.2)	20.6(1.7)	22.0(1.8)	6.2(1.1)	2.0(0.6)	0.4(0.3)	0.0(0.0)
			246.3(2.6)	231.5(4.2)	235.8(2.7)	246.6(8.1)	238.7(14.1)	236.5(17.0)	0.0(0.0)

Table 16-132

Weight Response Percentages and Proficiency Means (with Standard Errors) for Writing Main Samples, by Parents' Education Level, Grade 8

	N	WTD % (CV)	NOT HS	GRAD HS	POST HS	GRAD COL	UNKNOWN	MISSING
--TOTAL--	11086	100.0 (0%)	7.6(0.4)	24.5(0.8)	19.7(0.6)	40.3(1.1)	7.9(0.4)	0.2(0.0)
			250.5(1.7)	254.4(2.0)	264.6(1.0)	270.8(1.4)	240.8(2.2)	231.6(9.6)
GENDER								
MALE	5567	50.4 (1%)	6.2(0.3)	24.3(1.1)	18.9(0.6)	41.8(1.2)	8.8(0.6)	0.2(0.0)
			237.4(2.3)	243.3(2.3)	254.1(1.3)	261.5(1.5)	231.6(2.8)	230.2(11.8)
FEMALE	5519	49.6 (1%)	9.1(0.6)	24.8(0.8)	20.4(0.8)	38.7(1.2)	7.0(0.4)	0.1(0.1)
			259.6(2.7)	265.5(2.2)	274.4(1.5)	281.0(1.5)	252.5(2.8)	233.6(12.1)
DERIVED ETHNICITY/RACE								
WHITE	7071	69.8 (0%)	6.1(0.5)	24.3(1.1)	20.5(0.7)	43.5(1.4)	5.5(0.4)	0.1(0.0)
			255.6(2.4)	260.7(2.2)	269.3(1.3)	276.2(1.4)	246.3(2.6)	233.3(15.5)
BLACK	1778	15.5 (1%)	7.1(0.6)	28.2(1.3)	20.5(1.2)	33.7(1.3)	10.5(1.0)	0.2(0.1)
			240.8(4.6)	235.8(3.9)	248.7(2.7)	245.8(2.9)	231.5(4.2)	237.2(11.8)
HISPANIC	1616	10.2 (2%)	19.0(1.1)	25.0(1.2)	14.3(0.9)	24.5(1.6)	17.1(1.5)	0.2(0.1)
			245.6(3.1)	245.1(2.1)	254.5(4.7)	257.6(3.2)	235.8(2.7)	214.3(30.0)
ASIAN AMERICAN	420	2.7 (7%)	3.7(1.0)	7.9(1.5)	11.6(1.9)	58.3(4.7)	18.6(3.1)	0.2(0.2)
			262.4(14.1)	263.3(8.2)	259.8(8.4)	276.6(5.2)	246.6(8.1)	220.9(****)
AMERICAN INDIAN	163	1.4 (13%)	9.8(2.8)	23.2(4.1)	22.2(3.0)	33.8(4.7)	11.0(2.9)	0.8(0.8)
			229.1(16.1)	244.4(5.5)	256.5(5.7)	254.9(7.0)	238.7(14.1)	241.1(****)
PARENTAL EDUCATION								
LESS THAN HS	813	7.6 (5%)	100.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
			250.5(1.7)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
GRADUATED HS	2534	24.5 (3%)	0.0(0.0)	100.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
			0.0(0.0)	254.4(2.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
SOME EDU AFTER HS	2133	19.7 (3%)	0.0(0.0)	0.0(0.0)	100.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
			0.0(0.0)	0.0(0.0)	264.6(1.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
GRADUATED COLLEGE	4670	40.3 (3%)	0.0(0.0)	0.0(0.0)	0.0(0.0)	100.0(0.0)	0.0(0.0)	0.0(0.0)
			0.0(0.0)	0.0(0.0)	0.0(0.0)	270.8(1.4)	0.0(0.0)	0.0(0.0)
UNKNOWN	936	7.9 (5%)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	100.0(0.0)	0.0(0.0)
			0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	240.8(2.2)	0.0(0.0)

Table 16-133

Weight Response Percentages and Proficiency Means (with Standard Errors) for Writing Main Samples, by Gender, Grade 12

	N	WTD % (CV)	MALE	FEMALE	MISSING
--TOTAL--	11532	100.0 (1%)	48.9(0.6)	51.1(0.6)	0.0(0.0)
			275.3(1.2)	296.3(1.2)	0.0(0.0)
GENDER					
MALE	5515	48.9 (2%)	100.0(0.0)	0.0(0.0)	0.0(0.0)
			275.3(1.2)	0.0(0.0)	0.0(0.0)
FEMALE	6017	51.1 (1%)	0.0(0.0)	100.0(0.0)	0.0(0.0)
			0.0(0.0)	296.3(1.2)	0.0(0.0)
DERIVED ETHNICITY/RACE					
WHITE	7821	71.3 (1%)	49.1(0.8)	50.9(0.8)	0.0(0.0)
			279.9(1.2)	301.7(1.4)	0.0(0.0)
BLACK	1949	15.0 (3%)	46.3(1.4)	53.7(1.4)	0.0(0.0)
			257.3(2.1)	276.4(2.3)	0.0(0.0)
HISPANIC	1159	9.2 (5%)	49.6(2.8)	50.4(2.8)	0.0(0.0)
			266.3(2.8)	287.2(2.2)	0.0(0.0)
ASIAN AMERICAN	465	3.5 (5%)	51.8(2.6)	48.2(2.6)	0.0(0.0)
			280.9(4.0)	303.2(3.6)	0.0(0.0)
AMERICAN INDIAN	88	0.7 (16%)	58.5(6.8)	41.5(6.8)	0.0(0.0)
			260.4(8.0)	282.4(12.2)	0.0(0.0)
PARENTAL EDUCATION					
LESS THAN HS	875	7.8 (7%)	42.1(1.6)	57.9(1.6)	0.0(0.0)
			258.9(2.8)	279.8(2.2)	0.0(0.0)
GRADUATED HS	2420	22.1 (4%)	49.8(1.4)	50.2(1.4)	0.0(0.0)
			263.9(1.5)	285.6(1.8)	0.0(0.0)
SOME EDU AFTER HS	2993	25.9 (2%)	45.1(1.0)	54.9(1.0)	0.0(0.0)
			275.6(2.1)	296.5(1.9)	0.0(0.0)
GRADUATED COLLEGE	4952	41.7 (2%)	51.3(1.0)	48.7(1.0)	0.0(0.0)
			285.9(1.3)	307.4(1.5)	0.0(0.0)
UNKNOWN	265	2.2 (8%)	62.0(2.8)	38.0(2.8)	0.0(0.0)
			241.5(3.7)	264.5(4.7)	0.0(0.0)

Table 16-134

Weight Response Percentages and Proficiency Means (with Standard Errors) for Writing Main Samples, by Derived Race/ethnicity, Grade 12

	N	WTD % (CV)	WHITE	BLACK	HISPANIC	ASIAN	AMER IND	UNCLASSI	MISSING
--TOTAL--	11532	100.0 (1%)	71.3(0.5)	15.0(0.4)	9.2(0.4)	3.5(0.2)	0.7(0.1)	0.3(0.1)	0.0(0.0)
			291.0(1.1)	267.6(1.8)	276.9(1.8)	291.6(2.9)	269.5(8.2)	292.4(7.3)	0.0(0.0)
GENDER									
MALE	5515	48.9 (2%)	71.6(0.7)	14.3(0.6)	9.3(0.4)	3.7(0.3)	0.8(0.2)	0.3(0.1)	0.0(0.0)
			279.9(1.2)	257.3(2.1)	266.3(2.8)	280.9(4.0)	260.4(8.0)	282.3(12.2)	0.0(0.0)
FEMALE	6017	51.1 (1%)	71.0(0.7)	15.8(0.4)	9.0(0.8)	3.3(0.3)	0.6(0.1)	0.3(0.1)	0.0(0.0)
			301.7(1.4)	276.4(2.3)	287.2(2.2)	303.2(3.6)	282.4(12.2)	301.2(8.5)	0.0(0.0)
DERIVED ETHNICITY/RACE									
WHITE	7821	71.3 (1%)	100.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
			291.0(1.1)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
BLACK	1949	15.0 (3%)	0.0(0.0)	100.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
			0.0(0.0)	267.6(1.8)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
HISPANIC	1159	9.2 (5%)	0.0(0.0)	0.0(0.0)	100.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
			0.0(0.0)	0.0(0.0)	276.9(1.8)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
ASIAN AMERICAN	465	3.5 (5%)	0.0(0.0)	0.0(0.0)	0.0(0.0)	100.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
			0.0(0.0)	0.0(0.0)	0.0(0.0)	291.6(2.9)	0.0(0.0)	0.0(0.0)	0.0(0.0)
AMERICAN INDIAN	88	0.7 (16%)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	100.0(0.0)	0.0(0.0)	0.0(0.0)
			0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	269.5(8.2)	0.0(0.0)	0.0(0.0)
PARENTAL EDUCATION									
LESS THAN HS	875	7.8 (7%)	46.9(3.3)	16.8(2.2)	30.7(3.6)	3.5(0.9)	1.4(0.5)	0.6(0.4)	0.0(0.0)
			276.9(2.6)	255.1(3.5)	271.2(3.6)	266.3(10.3)	267.1(20.2)	286.3(14.6)	0.0(0.0)
GRADUATED HS	2420	22.1 (4%)	70.4(1.2)	18.4(1.1)	8.2(0.6)	2.0(0.3)	0.7(0.2)	0.3(0.1)	0.0(0.0)
			279.4(1.6)	258.1(2.0)	272.1(3.1)	276.0(5.4)	270.0(14.5)	289.0(14.3)	0.0(0.0)
SOME EDU AFTER HS	2993	25.9 (2%)	73.1(0.7)	17.0(0.7)	7.5(0.5)	1.8(0.2)	0.7(0.2)	0.0(0.0)	0.0(0.0)
			290.9(1.8)	272.6(2.8)	283.5(3.9)	289.3(8.9)	266.3(17.2)	298.8(33.8)	0.0(0.0)
GRADUATED COLLEGE	4952	41.7 (2%)	77.3(0.6)	11.3(0.5)	5.7(0.3)	4.8(0.3)	0.4(0.1)	0.5(0.1)	0.0(0.0)
			299.4(1.3)	278.0(2.7)	287.2(3.5)	302.2(3.5)	288.0(16.1)	299.5(8.2)	0.0(0.0)
UNKNOWN	265	2.2 (8%)	33.9(3.0)	23.0(2.9)	27.0(3.5)	11.5(3.2)	3.5(1.6)	1.1(1.2)	0.0(0.0)
			250.3(5.3)	238.1(8.1)	254.5(7.3)	267.0(6.5)	243.3(12.3)	249.6(5.2)	0.0(0.0)

BEST COPY AVAILABLE

Table 16-135

Weight Response Percentages and Proficiency Means (with Standard Errors) for Writing Main Samples, by Parents' Education Level, Grade 12

	N	WTD % (CV)	NOT HS	GRAD HS	POST HS	GRAD COL	UNKNOWN	MISSING
--TOTAL--	11505	100.0 (1%)	7.8(0.5)	22.2(0.7)	26.0(0.5)	41.8(1.0)	2.2(0.2)	0.3(0.1)
			271.0(1.9)	274.8(1.3)	287.1(1.7)	296.4(1.2)	250.3(3.1)	245.7(13.9)
GENDER								
MALE	5504	48.9 (2%)	6.7(0.5)	22.6(0.9)	24.0(0.6)	43.9(1.0)	2.8(0.3)	0.3(0.1)
			258.9(2.8)	263.9(1.5)	275.6(2.1)	285.9(1.3)	241.5(3.7)	239.4(16.3)
FEMALE	6001	51.1 (1%)	8.8(0.7)	21.8(0.8)	27.9(0.9)	39.9(1.3)	1.6(0.2)	0.3(0.1)
			279.8(2.2)	285.6(1.8)	296.5(1.9)	307.4(1.5)	264.5(4.7)	250.7(18.9)
DERIVED ETHNICITY/RACE								
WHITE	7808	71.3 (1%)	5.1(0.5)	21.9(0.8)	26.6(0.7)	45.3(1.2)	1.1(0.1)	0.2(0.1)
			276.9(2.6)	279.4(1.6)	290.9(1.8)	299.4(1.3)	250.3(5.3)	250.2(15.3)
BLACK	1943	15.0 (3%)	8.7(1.0)	27.1(1.5)	29.3(1.3)	31.5(1.2)	3.4(0.4)	0.3(0.1)
			255.1(3.5)	258.1(2.0)	272.6(2.8)	278.0(2.7)	238.1(8.1)	230.8(17.5)
HISPANIC	1153	9.1 (5%)	26.1(2.9)	20.0(1.4)	21.3(1.6)	26.0(2.0)	6.5(0.8)	0.6(0.3)
			271.2(3.6)	272.1(3.1)	283.5(3.9)	287.2(3.5)	254.5(7.3)	239.7(27.8)
ASIAN AMERICAN	465	3.5 (5%)	8.0(1.9)	13.1(2.1)	13.2(1.5)	58.5(4.5)	7.3(2.1)	0.0(0.0)
			266.3(10.3)	276.0(5.4)	289.3(8.9)	302.2(3.5)	267.0(6.5)	0.0(0.0)
AMERICAN INDIAN	86	0.7 (16%)	16.0(4.8)	23.5(3.8)	25.4(5.8)	23.8(5.5)	11.3(4.4)	2.8(2.2)
			267.1(20.2)	270.0(14.5)	266.3(17.2)	288.0(16.1)	243.3(12.3)	257.3(22.6)
PARENTAL EDUCATION								
LESS THAN HS	875	7.8 (7%)	100.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
			271.0(1.9)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
GRADUATED HS	2420	22.2 (4%)	0.0(0.0)	100.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
			0.0(0.0)	274.8(1.3)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
SOME EDU AFTER HS	2993	26.0 (2%)	0.0(0.0)	0.0(0.0)	100.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
			0.0(0.0)	0.0(0.0)	287.1(1.7)	0.0(0.0)	0.0(0.0)	0.0(0.0)
GRADUATED COLLEGE	4952	41.8 (2%)	0.0(0.0)	0.0(0.0)	0.0(0.0)	100.0(0.0)	0.0(0.0)	0.0(0.0)
			0.0(0.0)	0.0(0.0)	0.0(0.0)	296.4(1.2)	0.0(0.0)	0.0(0.0)
UNKNOWN	265	2.2 (8%)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	100.0(0.0)	0.0(0.0)
			0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	250.3(3.1)	0.0(0.0)

Table 16-136
Proficiencies for Writing Long-term Trend Samples, Grade 4

	1984	1988	1990	1992
-- TOTAL --	100.0(0.0) 203.8(1.5)	100.0(0.0) 205.7(1.6)	100.0(0.0) 201.7(1.5)*	100.0(0.0) 207.1(1.5)
GENDER				
MALE	53.3(1.7) 200.5(2.8)	50.4(1.0) 199.0(2.3)	51.2(1.0) 195.0(1.9)	50.4(0.6) 198.3(1.7)
FEMALE	46.7(1.7) 207.6(3.1)*	49.6(1.0) 212.6(2.0)	48.8(1.0) 208.7(2.2)*	49.6(0.6) 216.1(1.7)•
RACE/ETHNICITY				
WHITE	70.9(1.0) 210.7(1.9)	70.0(0.3) 214.9(1.9)	70.0(0.3) 211.0(2.0)	70.8(0.3) 216.7(1.7)•
BLACK	14.6(0.9) 181.6(5.0)	15.5(0.2) 173.3(4.7)	15.3(0.3) 171.4(5.4)	15.6(0.2) 175.0(3.8)
HISPANIC	11.4(0.7) 188.5(5.8)	10.6(0.3) 190.3(3.5)	10.8(0.2)* 184.1(4.1)	9.9(0.2)• 189.4(3.6)
OTHER	3.0(0.5) 208.2(6.1)	3.9(0.2) 211.0(4.8)	3.8(0.1) 201.7(7.0)	3.6(0.2) 206.9(2.6)
PARENTS' EDUCATION				
LESS THAN H.S.	6.6(0.6)* 178.7(4.6)	5.2(0.7) 194.2(5.4)	5.5(0.5) 185.7(3.9)	4.5(0.4)• 191.2(3.2)•
GRADUATED H.S.	20.5(1.1)* 191.8(3.4)	17.8(1.1) 199.4(3.0)	18.2(0.9) 196.6(3.0)	16.5(0.7)• 202.2(3.2)•
SOME EDUC AFTER H.S.	4.6(0.4) 208.0(6.5)	5.2(0.5) 211.4(6.3)	4.8(0.4) 213.9(4.0)	5.7(0.4)• 201.4(4.5)
GRADUATED COLLEGE	32.8(1.4)* 218.1(3.0)	41.3(1.5)▪ 212.4(2.2)	39.7(1.6)▪ 209.0(1.6)▪	41.8(1.0)• 213.7(1.4)
UNKNOWN	34.5(1.3) 202.8(2.5)	30.2(1.4) 201.5(2.7)	31.8(1.1) 196.5(2.2)*	31.1(1.0)• 204.7(2.3)

* Indicates a significant difference between 1992 and a previous year, at the 95% certainty level.

▪ Indicates a significant difference between the earliest year with available data and a subsequent year, at the 95% certainty level. These notations (*,▪) indicate statistical significance from a multiple comparison procedure based on the number of assessments with valid data.

• Indicates a significant difference between 1992 and the earliest year with availiable data, at the 95% certainty level. This notation indicates statistical significance without application of a multiple comparison procedure.

Table 16-137

Proficiencies for Writing Long-term Trend Samples, Grade 8

	1984	1988	1990	1992
-- TOTAL --	100.0(0.0) 266.7(2.0)*	100.0(0.0) 263.7(1.3)*	100.0(0.0) 256.6(1.2)▪*	100.0(0.0) 274.4(1.3)•
GENDER				
MALE	50.7(1.2) 257.5(2.3)	49.4(1.1) 253.7(1.5)*	50.5(0.9) 245.6(1.5)▪*	49.5(0.8) 263.6(1.9)•
FEMALE	49.3(1.2) 276.2(2.4)*	50.6(1.1) 273.5(1.7)*	49.5(0.9) 267.9(1.3)▪*	50.5(0.8) 285.0(1.3)•
RACE/ETHNICITY				
WHITE	75.7(0.9)* 271.7(2.1)*	70.6(0.2)▪ 269.1(1.3)*	70.2(0.2)▪ 262.1(1.6)▪*	70.3(0.3)• 279.2(1.3)•
BLACK	12.4(0.6)* 247.1(5.7)	14.9(0.2)▪* 246.0(3.5)	15.1(0.2)▪ 239.0(2.3)*	15.5(0.1)• 258.1(4.0)
HISPANIC	8.1(0.7)* 246.9(6.4)*	10.3(0.1)▪ 250.4(2.5)*	10.3(0.2)▪ 245.7(2.8)*	10.2(0.2)• 265.0(2.2)•
OTHER	3.8(0.4) 272.5(4.3)	4.2(0.1) 268.9(2.9)	4.4(0.1) 255 4(2.9)▪*	4.0(0.1) 275.5(4.8)
PARENTS' EDUCATION				
LESS THAN H.S.	10.0(0.8)* 257.7(4.8)	8.4(0.7) 254.3(3.9)	7.6(0.6) 245.6(3.7)	7.0(0.8)• 257.9(5.3)
GRADUATED H.S.	34.8(1.4)* 260.8(1.6)*	31.4(1.2) 257.6(2.1)*	32.7(1.1)* 252.5(1.4)▪*	28.8(1.1)• 267.9(1.6)•
SOME EDUC AFTER H.S.	9.5(0.8)* 271.4(3.9)	10.8(0.6) 275.1(3.3)	11.7(0.7) 266.9(3.0)*	12.4(0.7)• 279.6(2.2)
GRADUATED COLLEGE	35.8(1.5)* 277.8(1.8)*	41.0(1.5)▪ 270.5(1.8)▪*	38.1(1.5)* 264.9(1.8)▪*	43.9(1.8)• 284.3(1.9)•
UNKNOWN	8.9(0.8) 249.2(6.7)	8.0(0.5) 248.3(2.3)	9.6(0.6) 235.1(2.8)*	7.8(0.6) 249.6(3.1)

* Indicates a significant difference between 1992 and a previous year, at the 95% certainty level.

▪ Indicates a significant difference between the earliest year with available data and a subsequent year, at the 95% certainty level. These notations (*,▪) indicate statistical significance from a multiple comparison procedure based on the number of assessments with valid data.

• Indicates a significant difference between 1992 and the earliest year with availiable data, at the 95% certainty level. This notation indicates statistical significance without application of a multiple comparison procedure.

628

629

BEST COPY AVAILABLE

Table 16-138
Proficiencies for Writing Long-term Trend Samples, Grade 11

	1984	1988	1990	1992
-- TOTAL --	100.0(0.0) 289.7(1.6)	100.0(0.0) 291.3(1.3)	100.0(0.0) 287.1(1.0)	100.0(0.0) 287.3(1.4)
GENDER				
MALE	50.7(1.6) 281.1(1.4)	46.9(1.5)* 282.2(2.0)	50.9(0.8) 276.4(1.6)	53.1(1.1) 279.4(1.2)
FEMALE	49.3(1.6) 298.6(2.5)	53.1(1.5)* 299.3(1.2)	49.1(0.8) 298.2(1.5)	46.9(1.1) 296.4(2.0)
RACE/ETHNICITY				
WHITE	74.7(1.1)* 296.8(1.8)	73.7(0.1)* 296.2(1.3)	71.2(0.2)■ 292.8(1.2)	71.2(0.2)• 294.1(1.2)
BLACK	14.6(1.0) 270.3(3.6)	14.9(0.1)* 275.2(2.9)*	15.5(0.2)* 268.2(2.3)	13.9(0.2) 263.2(3.2)
HISPANIC	7.9(0.6)* 259.1(6.6)	8.0(0.1)* 273.8(4.4)	8.7(0.1)* 276.9(2.6)■	10.6(0.2)• 273.6(3.8)
OTHER	2.8(0.3)* 288.6(5.8)	3.4(0.1)■* 295.5(3.7)	4.5(0.0)■ 282.3(5.5)	4.4(0.2)• 287.1(3.0)
PARENTS' EDUCATION				
LESS THAN H.S.	10.7(1.2) 273.6(5.2)	8.2(0.8) 275.8(3.5)	8.5(0.5) 268.0(4.0)	8.2(0.8) 271.0(3.7)
GRADUATED H.S.	34.9(2.0)* 283.5(3.0)	30.4(1.2)* 284.6(2.2)	29.5(1.1) 278.2(1.9)	26.7(0.9)• 278.4(2.2)
SOME EDUC AFTER H.S.	14.9(0.9)* 297.6(2.5)	17.7(0.8) 296.1(2.6)	18.4(0.6)■ 292.2(2.7)	19.7(0.8)• 292.3(2.0)
GRADUATED COLLEGE	35.6(1.6)* 299.9(2.4)	40.9(1.8) 299.0(2.0)	40.4(1.4) 297.5(2.0)	42.6(1.4)• 295.7(1.4)
UNKNOWN	2.9(0.5) 261.3(7.3)	2.7(0.3) 267.6(4.7)	3.0(0.3) 259.0(4.1)	2.6(0.3) 259.1(6.6)

* Indicates a significant difference between 1992 and a previous year, at the 95% certainty level.

■ Indicates a significant difference between the earliest year with available data and a subsequent year, at the 95% certainty level. These notations (*,■) indicate statistical significance from a multiple comparison procedure based on the number of assessments with valid data.

• Indicates a significant difference between 1992 and the earliest year with availiable data, at the 95% certainty level. This notation indicates statistical significance without application of a multiple comparison procedure.

APPENDIX A

PARTICIPANTS IN THE OBJECTIVES AND ITEM DEVELOPMENT PROCESS

APPENDIX A
PARTICIPANTS IN THE OBJECTIVES AND ITEM DEVELOPMENT PROCESS

READING PROJECT STEERING COMMITTEE

Mike Cohen	National Governors' Association, Washington, DC
H. Dean Evans	Council of Chief State School Officers, Indianapolis, Indiana
Helene Hodges	Association of Supervision and Curriculum Development, Alexandria, Virginia
Kathleen Holliday	National Association of Elementary School Principals, Potomac, Maryland
Harriet C. Jelnek	National School Board Association, Rhineland, Wisconsin
Ann Kahn	National Parent Teacher Association, Alexandria, Virginia
Br. Robert Kealey	National Catholic Educational Association, Washington, DC
Paul Le Mahieu	National Association of Test Directors, Pittsburgh, Pensylvania
Gary Marx	American Association of School Administrators, Arlington, Virginia
Carole Perlman	American Educational Research Association, Chicago, Illinois
Marily Rauth	American Federation of Teachers, Washington, DC
Edward Roeber	Association of State Assessment Programs, Lansing, Michigan
Esther Schaeffer	National Alliance of Business, Washington, DC
Ann Smith	National Educational Association, Ormond Beach, Florida
Scott Thompson	National Education of Secondary School Principals, Reston, Virginia

READING PROJECT PLANNING COMMITTEE

Marilyn Adams	BBN Laboratories, Cambridge, Massachusetts
Marsha Delain	South Carolina Department of Education, Columbia, South Carolina
Lisa Delpit	Institute for Urban Research, Morgan State University, Baltimore, Maryland
William Feehan	Chase Manhattan Bank, New York, New York
Philip Gough	Department of Psychology, University of Texas at Austin, Austin, Texas
Edward Haertel	Stanford University, Stanford, California
Elfrieda Hiebert	School of Education, University of Colorado, Boulder, Colorado
Janet Jones	Charles County Public Schools, Waldorf, Maryland

Judith Langer	School of Education, State University of New York, Albany, Albany, New York
P. David Pearson	University of Illinois College of Education, Champaign, Illinois
Charles Peters	Oakland Schools, Pontiac, Michigan
John P. Pikulski	College of Education, University of Delaware, Newark, Delaware
Keith Stanovich	Oakland University, Rochester, Michigan
Paul Randy Walker	Maine Department of Education, Augusta, Maine
Sheila Valencia	University of Washington, Seattle, Washington

1992 NAEP READING CCSSO PROJECT STAFF

Ramsay W. Selden	Director, State Education Assessment Center, Council of Chief State School Officers
Barbara Kapinus	Project Coordinator, Council of State School Officers
Diane Schilder	Project Associate, Council of State School Officers

THE READING ITEM DEVELOPMENT COMMITTEE

Mary Barr	San Diego, CA
Carita Chapman	Chicago, IL
Richard Halle	Marshfield, WI
Elfrieda Hebert	School of Education, University of Colorado, Boulder, CO
Barbara Kapinus	Hyattsville, MD
Judith Langer	School of Education, SUNY - Albany, Albany, NY
Edye Norniella	Miami, FL
Charles Peters	Oakland Schools, Waterford, MI
John Pikulski	Newark, DE
Robert Swartz	Newtonville, MA

NATIONAL ASSESSMENT MATHEMATICS PLANNING PROJECT

Mathematics Steering Committee

Robert Astrup	National Education Association
Lillian Barna	Council of the Great City Schools
Richard A. Boyd	Council of Chief State School Officers

Glenn Bracht	Council for American Private Education and National Association of Independent Schools
William M. Ciliate	National School Boards Association
Antonia Cortese	American Federation of Teachers
Mary Brian Costello	National Community on Catholic Education Association
Wilhelmina Delco	National Council of State Legislators
Nancy DiLaura	National Governors' Association
Thomas Fisher	Association of State Assessment Programs
Alice Houston	Association for Supervision and Curriculum Development
C. June Knight	National Association of Elementary School Principals
Stephen Lee	National Association of Secondary School Principals
Paul LeMahieu	National Association of Test Directors
Glen Ligon	Directors of Research and Evaluation
Barbara Roberts Mason	National Association of State Boards of Education
James E. Morrell	American Association of School Administrators, Austin Independent School District, Texas

Mathematics Objectives Committee

Joan Burks	Damascus High School, Damascus, Maryland
Phillip Curtis	University of California at Los Angeles, Los Angeles, California
Walter Denham	California Department of Education, Sacramento, California
Thomas Fisher	Florida Department of Education, Tallahassee, Florida
Ann Kahn	The National Parent-Teacher Association, Fairfax, Virginia
Mary M. Lindquist	Columbus College, Columbus, Georgia
Susan Purser	Whitten Junior High School, Jackson, Mississippi
Dorothy Strong	Chicago Public Schools, Chicago, Illinois
Thomas W. Tucker	Colgate University, Hamilton, New York
Charles Watson	Arkansas Department of Education, Little Rock, Arkansas
O.R. Wells, Jr.	Rice University, Houston, Texas

NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS

Mathematics Item Development Panel

Martha Baca	Roosevelt School District, Phoenix, Arizona

Iris Carl	Houston Independent School District, Houston, Texas
Carol Cho	Alhambra High School, Martinez, California
John Dossey	Illinois State University, Normal, Illinois
Elizabeth Fennema	University of Wisconsin, Madison, Wisconsin
Steven Leinwand	Connecticut State Department of Education, Middletown, Connecticut
Mary M. Lindquist	Columbus College, Columbus, Georgia
Gloria Sanok	Packanack School, West Caldwell, New Jersey
Thomas Tucker	Colgate University, Hamilton, New York

Mathematics Test Development Consultants

James Braswell	College Board Programs, Educational Testing Service
Gail Chapman	College Board Programs, Educational Testing Service
Jeanne Galick	College Board Programs, Educational Testing Service
Jeffrey Haberstroh	College Board Programs, Educational Testing Service
Chancey Jones	College Board Programs, Educational Testing Service
Patricia Klag	College Board Programs, Educational Testing Service
Jane Kupin	College Board Programs, Educational Testing Service
Jane Maroney	College Board Programs, Educational Testing Service
Marlene Supernavage	College Board Programs, Educational Testing Service
Beverly Whittington	College Board Programs, Educational Testing Service

Mathematics Scale Anchoring Panel

Laurie Boswell	Profile High School, Bethlehem, New Hampshire
Bruce Brombacher	Jones Junior High School, Westerville, Ohio
Catherine Brown	Virginia Polytechnic Institute, Blacksburg, Virginia
Joe Crosswhite	Springfield, Missouri
John Dossey	Illinois State University, Normal, Illinois
Henry Kepner, Jr.	University of Wisconsin at Milwaukee, Milwaukee, Wisconsin
Linda Kolnowski	Detroit Public Schools, Detroit, Michigan
Gordon Lewis	Washington DC Public Schools, Washington, D.C.
Mary M. Lindquist	Columbus College, Columbus, Georgia

Donna Long	Indiana Dept. of Education, Indianapolis, Indiana
Vena Long	Missouri Department of Education, Jefferson City, Missouri
William Masalski	University of Massachusetts, Amherst, Massachusetts
Wendell Meeks	Illinois State Board of Education, Springfield, Illinois
Andy Reeves	Florida Department of Education, Tallahassee, Florida
Diane Thiessen	University of Northern Iowa, Cedar Falls, Iowa
Alba Thompson	CRMSE, San Diego, CA 92120
Shiela Vice	Kentucky Department of Education, Frankfort, Kentucky
Charles Watson	Arkansas Department of Education, Little Rock, Arkansas
Vernon Williams	H.W. Longfellow Intermediate School, Falls Church, Virginia

WRITING DEVELOPMENT PARTICIPANTS

Benita Albert	Oakridge, Tennessee
Arthur Applebee	State University of New York at Albany, Albany, New York
Katherine Au	Kamehameha School, Honolulu, Hawaii
Martha Baca	Phoenix, Arizona
Elizabeth Badger	Quincy, Massachusetts
Linda Bailey	Oklahoma City, Oklahoma
Jeffrey Barnett	University of Wisconsin at Whitewater, Whitewater, Wisconsin
Mary Barr	California Literature Project, San Diego, California
Pat Bates	Timber Lane Junior High School, Pennington, New Jersey
John Beagan	Princeton University, Princeton, New Jersey
Michael Beck	Beta, Inc., Pleasantville, New York
Mary Bedard	Trenton, New Jersey
Brigitte Bentele	New York City, New York
Lucy Bikulcs	Chicago, Illinois
Marilyn Bloch	Montgomery, Alabama
Barbara Braverman	West Windsor-Plainsboro Middle School, Plainsboro, New Jersey
Diane Briars	Pittsburgh Public Schools, Pittsburgh, Pennsylvania
Nancy Broz	Haddonfield, New Jersey
Thomas Butts	Dallas, Texas
Norma Byers	Stuart Country Day School, Princeton, New Jersey

Candy Carter	Truckee, California
Vincent Casessa	Hoboken, New Jersey
Edward Chango	J. R. Masterman School, Philadelphia, Pennsylvania
Carita Chapman	Chicago, Illinois
Peg Cibes	New London, Connecticut
Terrence Coburn	Oakland Schools, Pontiac, Michigan
Kathleen Connell	Iowa City, Iowa
Ann Conway	Langhorne Manor, Pennsylvania
Thomas Cooney	University of Georgia, Athens, Georgia
Kathy Connell	Iowa City, Iowa
Marlene Crane	Williamsburg, Missouri
Cheryl David	St. Thomas, Virgin Islands
Joseph Dodson	Winston-Salem/Forsyth County Schools, Winston-Salem, North Carolina
Carolyn Ehr	Fort Hays State University, Hays, Kansas
Pamela Elvington	Pennington, New Jersey
Elizabeth Fennema	University of Wisconsin at Madison, Madison, Wisconsin
John Finch	Park Hills, Kentucky
Lynn Fleck	Yukon, Oklahoma
Johnetta Fleming	Los Angeles, California
Dan Foertsch	Lawrenceville, New Jersey
Linda Foreman	Oregon City, Oregon
Sarah Freedman	University of California at Berkeley, Berkeley, California
Elizabeth Fuhrman	Rutgers University, New Brunswick, New Jersey
Lucy Garner	Los Angeles Center for Enriched Studies, Los Angeles, California
Lorraine Gerhart	Oconomowoc, Wisconsin
David Glatzer	West Patterson, New Jersey
Joyce Glatzer	West Patterson, New Jersey
Heriberto Godina	El Paso, Texas
Myron Goldman	J. R. Masterman School, Philadelphia, Pennsylvania
Frank Gunnip	Ferndale, Michigan
James Guthrie	University of California at Berkeley, Berkeley, California

Richard Halle	Marshfield, Wisconsin
Gary Havenstein	Tracey, Indiana
James Heiman	Blue Springs, Missouri
Geoff Hewitt	Vermont State Department of Education, Montpelier, Vermont
Elfrieda Hiebert	University of Colorado, Boulder, Colorado
Connie Hora	Iowa City, Iowa
Nola Hoss	Danielson, Connecticut
Shelly Hughes	Middleburg, Florida
Jeanne Joyner	North Carolina Department of Public Instruction, Raleigh, North Carolina
Barbara Kapinus	Maryland State Department of Education, Baltimore, Maryland
Patrick Kent	Westerville, Ohio
Mary Kitagawa	Amherst, Massachusetts
Vicky Kouba	State University of New York, Albany, New York
Judith Langer	State University of New York, Albany, New York
Annita Larson	Lansing, Michigan
Shawn Lee	Iowa City, Iowa
Regina Lemerick	Princeton Day School, Princeton, New Jersey
Reynaldo Macias	University of Southern California, Alhambra, California
Nancy Mack	Northern Illinois University, DeKalb, Illinois
Mary Meng	Deptford, New Jersey
Paul Moraski	Gouldsboro, Pennsylvania
Anita Morrisey	Gardiner, Maine
Marilyn Nance	Owasso, Oklahoma
Rochelle Newman	West Windsor-Plainsboro Middle School, Plainsboro, New Jersey
Edith Norniella	Miami, Florida
Phares O'Daffer	Hudson, Illinois
David Pagni	California State University at Fullerton, Fullerton, California
Kathy Peal	Pendleton, Oregon
Barbara Pence	San Jose State University, San Jose, California
Charles Peters	Oakland Schools, Pontiac, Michigan
Terry Peterson	Winthrop College, Columbia, South Carolina

Sherry Phinney	Antrim, New Hampshire
John Pikulski	University of Delaware, Newark, Delaware
Sheldon Post	Richboro Junior High School, Richboro, Pennsylvania
Senta Raizen	National Center for Improving Science Education, Washington, DC
Claudia Runge	Louisville, Kentucky
Joseph Santora	Essex County College, Newark, New Jersey
Harriet Sawyer	Northville, Michigan
Chuck Schaub	Pennington, New Jersey
Jane Seeburger	Skillman, New Jersey
Michael Shaughnessy	Oregon State University, Corvallis, Oregon
Jack Smith	Hermitage, Pennsylvania
Phyllis Stern	Edgewater Park, New Jersey
Marlene Supernavage	Lambertville, New Jersey
Robert Swartz	Newtonville, Massachusetts
William Taylor	East Lyme, Connecticut
Linda Tetley	Jefferson City, Missouri
Laura Tsosie	Pinon Public School, Pinon, Arizona
Alba Thompson	Illinois State University, Normal, Illinois
Terry Underwood	Citrus Heights, California
Jolean Wegner	Waukesha, Wisconsin
Ellen Wehrman	Bordentown, New Jersey
Sharon Weinman	Gardendale, Alabama
Norma Wiesen	East Hardwick, Vermont
Clyde Wiles	Indiana University Northwest, Gary, Indiana
Linda Winfield	Johns Hopkins University, Baltimore, Maryland
Jan Winston	Coffeyville, Kansas
Les Winters	Northridge, California
Shirley Wiseman	Marion, Indiana
Willy Wood	Middlebury, Vermont
Dottie Woodbury	Stuart Country Day School, Princeton, New Jersey
Rosemary Zurawel	Dover, New Hampshire

APPENDIX B

NAEP REPORTING SUBGROUPS

COMPOSITE AND DERIVED REPORTING AND CONDITIONING VARIABLES

REPORTING SUBGROUPS FOR THE 1992 NATIONAL ASSESSMENT

Results for the 1992 National Assessment were reported for student subgroups defined by gender, race/ethnicity, type of community, parents' level of education, and geographical region. The following explains how each of these subgroups was derived.

DSEX (Gender)

The variable SEX is the gender of the student being assessed, as taken from school records. For a few students, data for this variable was missing and was imputed by ETS after the assessment. The resulting variable DSEX contains a value for every student and is used for gender comparisons among students.

DRACE (Self-identified Race/ethnicity)

The variable DRACE is an imputed definition of race/ethnicity, derived primarily from student self-identification. This variable is used for race/ethnicity subgroup comparisons in the 1992 main samples and in the 1992 writing trend samples (see also "Observed race/ethnicity" below). Two items from the student demographics questionnaire were used in the determination of self-identified race/ethnicity:

Demographic Item Number 2:

2. If you are Hispanic, what is your Hispanic background?

○ I am not Hispanic.
○ Mexican, Mexican American, or Chicano
○ Puerto Rican
○ Cuban
○ Other Spanish or Hispanic background

Students who responded to item number 2 by filling in the second, third, fourth, or fifth oval were considered Hispanic. For students who filled in the first oval, did not respond to the item, or provided information that was illegible or could not be classified, responses to item number 1 were examined in an effort to determine race/ethnicity. Item number 1 read as follows:

Demographic Item Number 1:

1. Which best describes you?

○ White (not Hispanic)

○ Black (not Hispanic)

○ Hispanic ("Hispanic" means someone who is Mexican, Mexican American, Chicano, Puerto Rican, Cuban, or from some other Spanish or Hispanic background.)

○ Asian or Pacific Islander ("Asian or Pacific Islander" means someone who is Chinese, Japanese, Korean, Filipino, Vietnamese, or from some other Asian or Pacific Island background.)

○ American Indian or Alaskan Native ("American Indian or Alaskan Native" means someone who is from one of the American Indian tribes, or one of the original people of Alaska.)

○ Other (What?) ________

Students' race/ethnicity was then assigned to correspond with their selection. For students who filled in the sixth oval ("Other"), provided illegible information or information that could not be classified, or did not respond at all, race/ethnicity as provided from school records was used.

Race/ethnicity could not be determined for students who did not respond to background items 1 or 2 and for whom race/ethnicity was not provided by the school.

RACE (Observed race/ethnicity)

The variable RACE is the race/ethnicity of the student being assessed as observed and recorded by the exercise administrator. Observed race/ethnicity was used in NAEP assessments before 1984 (see also "Self-identified race/ethnicity" above). Observed race/ethnicity is used for all race/ethnicity subgroup trend comparisons for which the starting trend point comes from pre-1984 assessment data. Table B-1 shows which of the race/ethnicity variables (observed or self-identified) were used for each subject area report for the 1992 assessment.

Table B-1
Race/ethnicity Definitions Used in NAEP Reports for the 1992 Assessment

Report	Race/ethnicity Definition Used
NAEP 1992 Reading Report Card	Self-identified
NAEP 1992 Mathematics Report Card	Self-identified
NAEP 1992 Writing Report Card	Self-identified
NAEP 1992 Trends in Academic Progress: Achievement of U.S. Students in Science, 1969 to 1992; Mathematics, 1973 to 1992; Reading, 1971 to 1992; Writing, 1984 to 1992	Self-identified for writing; observed for reading, mathematics, and science
NAEP's 1992 Writing Portfolio Study: Refining Methods for Collecting and Evaluating Students' School-based Writing	Self-identified

TOC (Type of community)

NAEP assigned each participating school to one of four type of categories designed to provide information about the communities in which the schools are located.

The type of community categories consist of three "extreme" types of communities and one "other" type of community. Schools were placed into these categories on the basis of information about the type of community, the size of its population (as of the 1980 Census), and an occupational profile of residents provided by school principals before the assessment. The principals completed estimates of the percentage of students whose parents fit into each of six occupational categories. For those schools where the principal or his or her designate was unable or unwilling to answer the question on the occupational profile of parents, the type of community category was assigned as "missing." The type of community categories are listed and described below.

> 1 - Extreme Rural: Students in this group live outside metropolitan statistical areas, live in areas with a population below 10,000, and attend schools where many of the students' parents are farmers or farm workers.
>
> 2 - Disadvantaged Urban: Students in this group live in metropolitan statistical areas and attend schools where a high proportion of the students' parents are on welfare or are not regularly employed.
>
> 3 - Advantaged Urban: Students in this group live in metropolitan statistical areas and attend schools where a high proportion of the students' parents are in professional or managerial positions.
>
> 4 - Other: Students in this category attend schools in areas other than those defined as advantaged urban, disadvantaged urban, or extreme rural.

Extreme Rural. Schools were assigned a rural index based on occupation percentages and size of community from the principal's questionnaire. Only school meeting the following conditions were eligible for consideration as extreme rural: The school had to be in a place outside a metropolitan statistical area; the percentage of farm workers had to be nonzero; and the size-of-community percentages on the principal's questionnaire had to be nonzero for rural areas and zero for towns in excess of 10,000.

For these eligible schools, the rural index was calculated by the formula

$$P_1 - (P_2 + 2P_3)$$

where

P_1 = percentage of farm workers;
P_2 = percentage of factory workers; and
P_3 = percentage of professionals or managers.

High values of this index result from relatively high percentages of persons employed in agriculture and relatively low percentages in professional, managerial, and blue collar jobs. Schools not qualifying for consideration as extreme rural were assigned a dummy index of (-200), which would place them at the low end of a descending order of the rural index. The first 10 percent of the schools in the described ranking were assigned to this category. The cumulation to 10 percent was based on the weighted student samples, so that 10 percent of the weighted assessed students were in schools assigned to the extreme rural category.

Disadvantaged Urban. The same general approach used for the extreme rural category was used to identify disadvantaged urban schools. Only schools in metropolitan statistical areas were eligible. For these schools, an index was computed with the formula

$$P_4 + P_5 - P_3)$$

where

P_4 = percentage unemployed;
P_5 = percentage on welfare; and
P_3 = percentage of professionals or managers.

High values of the index were obtained for schools with relatively high percentages of unemployed and on welfare and relatively low percentages of professional and managerial occupations. Noneligible schools were assigned a dummy index of (-200). Schools were then ranked in descending order of the index. The first 10 percent of the schools in the ordered list were assigned to the disadvantaged urban category. The cutoffs were such that 10 percent of the weighted assessed students belonged to the disadvantaged urban category.

Advantaged Urban. This category was also restricted to schools in metropolitan statistical areas. An index was computed using the formula

$$P_3 - (P_1 + P_2 + P_4 + P_5)$$

where the terms are defined as for the extreme rural and disadvantaged urban categories. High values of this index were obtained for schools with relatively high percentages of professional and managerial personnel and relatively low percentages of blue collar workers, agricultural workers, unemployed persons, and welfare recipients. Noneligible schools were assigned a dummy index of (-200). Schools were then ranked in descending order of the index. The first 10 percent of the schools in this ranking were assigned to the advantaged urban category.

Other. This category was assigned to all schools not included in one of the other categories.

PARED (Parents' education level)

The variable PARED is derived from responses to two questions, B003501 and B003601, in the student demographic questionnaire. Students were asked to indicate the extent of their mother's education (B003501—How far in high school did your mother go?) by choosing one of the following:

○ She did not finish high school.
○ She graduated from high school.
○ She had some education after high school.
○ She graduated from college.
○ I don't know.

Students were asked to provide the same information about the extent of their father's education (B003601—How far in high school did your father go?) by choosing one of the following:

○ He did not finish high school.
○ He graduated from high school.
○ He had some education after high school.
○ He graduated from college.
○ I don't know.

The information was combined into one parental education reporting category (PARED) as follows: If a student indicated the extent of education for only one parent, that level was included in the data. If a student indicated the extent of education for both parents, the higher of the two levels was included in the data. For students who did not know the level of education for both parents or did not know the level of education for one parent and did not respond for the other, the parental education level was classified as unknown. If the student did not respond for both parents, the student was recorded as having provided no response.

REGION (Region of the country)

States were grouped into four geographical regions—Northeast, Southeast, Central, and West—as shown in Table B-2. All 50 states and the District of Columbia are listed, with the participants in the Trial State Assessment shown in italic type. Territories were not assigned to a region. The part of Virginia that is included in the Washington, DC, metropolitan statistical area is included in the Northeast region; the remainder of the state is included in the Southeast region.

Table B-2
NAEP Geographic Regions

NORTHEAST	SOUTHEAST	CENTRAL	WEST
Connecticut	*Alabama*	Illinois	Alaska
Delaware	*Arkansas*	*Indiana*	*Arizona*
District of Columbia	*Florida*	*Iowa*	*California*
Maine	*Georgia*	Kansas	*Colorado*
Maryland	*Kentucky*	*Michigan*	*Hawaii*
Massachusetts	*Louisiana*	*Minnesota*	*Idaho*
New Hampshire	*Mississippi*	*Missouri*	Montana
New Jersey	*North Carolina*	*Nebraska*	Nevada
New York	*South Carolina*	*North Dakota*	*New Mexico*
Pennsylvania	*Tennessee*	*Ohio*	*Oklahoma*
Rhode Island	*Virginia*	South Dakota	Oregon
Vermont	*West Virginia*	*Wisconsin*	*Texas*
Virginia			*Utah*
			Washington
			Wyoming

DGRADE, MODGRD (Grade in school)

To enhance the usefulness of the data, in 1984 NAEP began sampling students by grade as well as by age. The ages sampled in assessments since 1984——9, 13, and 17——match the ages sampled in earlier assessments. However, some of the modal grades (the grade attended by most students of a particular age) for the ages sampled have varied in the last three assessments because of changes in how student age was determined and changes in the times of the year that students were tested.

In the 1992 main assessment, the respective modal grades for ages 9, 13, and 17 are 4, 8, and 12. Student age for all three cohorts was determined on a calendar-year basis; all students were tested at the same times of the year. The 1990 long-term trend samples, by definition designed to match previous assessment characteristics, sampled varying student cohorts.

The 1992 main sample included many students in each cohort who were both age-eligible (age 9, 13, or 17) and grade-eligible (attending respectively grade 4, 8, or 12). However, because NAEP collected data by grade or age, each cohort also included students who were age-eligible but not in the modal grade, and students who were grade-eligible but not of the modal age (the age of most students attending the particular grade).

For each 1992 sample, results for students in a particular grade can be selected using (1) the variable DGRADE, the student's actual grade at time of testing, on the student file, or (2) the student file variable MODGRD (setting MODGRD to a value of 2 will select those students who are in the modal grade).

DAGE, MODAGE (Student age)

Results for students at a particular age can be selected using (1) the student file variable DAGE, the student's age as of December 31, 1991 or (2) the student file variable MODAGE. Because NAEP collected data by grade *or* age, each main sample student cohort includes students who were both age-eligible and grade-eligible, students who were age-eligible but not in the modal grade, and students who were grade-eligible but not of the modal age. The main assessment modal age (the age of most of the students in the grade sample) is age 9 for fourth graders, age 13 for eighth graders, and age 17 for twelfth graders. A value of 1 for MODAGE indicates that the student is younger than the modal age; a value of 2 indicates that the student at the modal age; a value of 3 indicates that the student is older than the modal age.

DERIVED VARIABLES USED FOR REPORTING AND CONDITIONING

Several NAEP variables were formed for use in reporting and conditioning from the systematic combination of response values for one or more items from either the student demographic questionnaire, the student subject-area background questionnaires, the teacher questionnaire, or the school questionnaire.

These variables maximize use of the data, incorporate a larger segment of the population, and save analysis costs by grouping items that measure similar characteristics into one variable. The derivation of each of these variables is explained below.

Variables Derived from the Student Demographic Questionnaires

HOMEEN2 (Home environment——Articles [of 4] in the home)

The variable HOMEEN2 was created from the responses to student demographic items B000901 (Does your family get a newspaper regularly?), B000903 (Is there an encyclopedia in your home?), B000904 (Are there more than 25 books in your home?), and B000905 (Does your family get any magazines regularly?). The values for this variable were derived as follows:

1 0-2 types	The student responded to at least two items and answered Yes to two or fewer.

2	3 types	The student answered Yes to three items.
3	4 types	The student answered Yes to four items.
8	Omitted	The student answered fewer than two items.

PARWK4 (Economic Support—Which parents work)

For age class 9, PARWK4 was created from responses to items B005901 and B006101, which asked if the student's mother (or stepmother) and father (or stepfather) worked for pay. The values for PARWK4 were derived as follows:

1	Both mother & father	The student answered YES to both B005901 and B006101.
2	Father only	The student answered NO or DON'T LIVE WITH to B005901 and YES to B006101.
3	Mother only	The student answered NO or DON'T LIVE WITH to B006101 and YES to B005901.
4	Something else	Any other combination of responses
8	No response	The student did not respond to one or both items.
9	Mult.	The student filled in more than one oval for both items

NCOMP (Number of computer science courses taken)

For age class 17, NCOMP was created from responses to items B005312 and B005313 concerning the student's coursework in computer science. The values for NCOMP were derived as follows:

1	0	The student answered HAVE NOT to both courses.
2	1	The student answered HAVE to one course.
3	2	The student answered YES to both courses.
8	No response	The student did not respond to one or both items.
9	Mult. & out-of-range	The student filled in more than one oval for both items.

NMATH (Highest level of mathematics courses taken)

For age class 17, NMATH was created from responses to items B005301 through B005307 concerning the student's coursework in mathematics. The values for NMATH were derived as follows:

1	Gen. math or pre-algebra	The student answered HAVE NOT to all items or HAVE to B005301 or B005302 and HAVE NOT to all others.
2	Algebra	The student answered HAVE to B005303 and HAVE NOT to B005304, B005305, B005306, and B005307.
3	Geometry	The student answered HAVE to both B005303 and B005305 and HAVE NOT to B005304, B005306, and B005307.
4	Algebra 2	The student answered HAVE to B005304 or B005306 but HAVE NOT to B005307.
5	Calculus	The student answered HAVE to B005307.
6	Something else	Any other response combination
8	No response	The student did not respond to any item.

NSCI (Highest level of science courses taken)

For age class 17, NSCI was created from responses to items B005308 through B005311, which concerned the student's coursework in science. The values for NSCI were derived as follows:

1	No biology	The student answered HAVE NOT to all items or HAVE to B005308 and other than HAVE to all other items.
2	Biology	The student answered HAVE to B005309 and other than HAVE to both B005310 and B005311.
3	Chemistry	The student answered HAVE to both B005309 and B005310 and other than HAVE to B005311.
4	Physics	The student answered HAVE to B005309, B005310, and B005311.
5	Something else	Any other response combination
8	No response	The student answered none of the items.

SINGLEP (How many parents live at home)

SINGLEP was created from items B005601 (Does either your mother or your stepmother live at home with you?) and B005701 (Does either your father or your stepfather live at home with you?). The values for SINGLEP were derived as follows:

1	2 parents at home	The student answered Yes to both items.
2	1 parent at home	The student answered Yes to B005601 and No to B005701, or Yes to B005701 and No to B005601.
3	Neither at home	The student answered No to both items.
8	Omitted	The student did not respond to or filled in more than one oval for one or both items.

Variables Derived from the Mathematics Background Questionnaires

ALGCALC (Algebra and calculus course taking, age 17/grade 12)

ALGCALC was created from five of the items in the mathematics background questionnaire that asked students how long they had taken certain mathematics courses:

M811003	Introduction to algebra or pre-algebra
M811004	First-year algebra
M811006	Second-year algebra
M811008	Pre-calculus, third-year algebra, elementary functions, or analysis
M811011	Calculus

For each item, the student could respond as follows:

1. More than 1 year
2. 1 school year
3. ½ year or less
4. Not studied

The values for ALGCALC were derived as follows:

Not studied	The student did not answer MORE THAN 1 YEAR or 1 SCHOOL YEAR to M811003 (half-year courses were counted as not studied)
Pre-algebra	The student answered MORE THAN 1 YEAR or 1 SCHOOL YEAR to M811003 but not to M811004
1st year algebra	The student answered MORE THAN 1 YEAR or 1 SCHOOL YEAR to M811004 but not to M811006

2nd year algebra	The student answered MORE THAN 1 YEAR or 1 SCHOOL YEAR to M811006 but not to M811008
3rd year algebra	The student answered MORE THAN 1 YEAR or 1 SCHOOL YEAR to M811008 but not to M811011
Calculus	The student answered MORE THAN 1 YEAR or 1 SCHOOL YEAR to M811011

GEOTRIG (Geometry and trigonometry course taking, age 17/grade 12)

GEOTRIG was created from two of the items in the mathematics background questionnaire that asked students how long they had taken certain mathematics courses:

M811005	Geometry
M811007	Trigonometry

For each item, the student could respond as follows:

1. More than 1 year
2. 1 school year
3. ½ year or less
4. Not studied

The values for GEOTRIG were derived as follows:

Not studied	The student did not answer MORE THAN 1 YEAR or 1 SCHOOL YEAR to M811005 (half-year courses were counted as not studied)
Geometry	The student answered MORE THAN 1 YEAR or 1 SCHOOL YEAR to M811005 but not to M811007
Trigonometry	The student answered MORE THAN 1 YEAR or 1 SCHOOL YEAR to M811007

Variables Derived from the Writing Background Questionnaires

WSENWP (Students' reports on teachers' encouragement of the writing process)

Eighth- and twelfth-grade students' responses to W801204 (How often are you asked by your teacher to plan writing?) and W801202 (How often are you asked by your teacher to write more than 1 draft?) were combined to form WSENWP. The number of times a student responded "Always" to each of these questions were summed. The range, therefore, is 1 to 3, where 1 means that planning and writing more than one draft is not always asked, and 3 means that both planning and writing more than one draft are always done. If a student did not answer either item, variable WSENWP was treated as omitted for that student.

Variables Derived from the Reading Teacher Questionnaire

TRCERTO (Teaching certificate - Other)

The variable TRCERTO was created from the responses to teacher background questions T040504 (Do you have teaching certification in middle/junior high school or secondary mathematics?) and T040505 (Do you have teaching certification in [some] other [category]?). The values for this variable were defined as follows:

1	Yes	Teacher indicated they were certified in mathematics or they were certified in "other."
2	No	Teacher indicated that they were not certified in either mathematics or "other."
3	Not offered	Teacher indicated that neither mathematics nor "other" certification was offered in their state.

TRUMAJB (Teacher undergraduate major - Other)

This variable was based on teachers' indications as to whether they had an undergraduate major in mathematics (T040703), mathematics education (T040704), or some other area (T040705).

1	Yes	Undergraduate major in mathematics, mathematics education, or "other."
2	No	No undergraduate major in mathematics, mathematics education, or "other."
8	Omitted	Teacher did not provide responses to any of the undergraduate majors.

TRGMAJB (Teacher graduate major - Other)

This variable was based on teachers' indications as to whether they had a graduate major in mathematics (T040803), mathematics education (T040804), or some other area (T040805).

1	Yes	Graduate major in mathematics, mathematics education, or "other."
2	No	No graduate major in mathematics, mathematics education, or "other."
8	Omitted	Teacher did not provide responses to any of the graduate majors.

TRTRAIN (Teacher training in reading activities)

This variable was created by examining teachers' responses to whether they had any training, in either college courses or in-service education, in eight areas:

T045903	Teaching critical thinking
T045904	The role of students' prior knowledge in their reading
T045907	Literature-based reading instruction
T045908	Reading assessment
T045909	Content area reading
T045910	Combining reading and writing
T045911	The whole language approach to teaching reading
T045912	Phonics in the teaching of reading

The number of times the teacher said "yes" to these questions was summed. The values of TRTRAIN were assigned as follows.

1	0-3	0 to 3 areas of training
2	4-6	4 to 6 areas of training
3	7	7 areas of training
4	8	8 areas of training

TREMP1 (Teaching heavy emphasis #1)

The responses to T046801 (How much instructional time is devoted to decoding skills?) and T046901 (How much emphasis is given to phonics?) were used to create TREMP1. If the teacher responded "almost all of the time" for T046801 and "heavy emphasis" for T046901, then TREMP1 was given a value of 1. Other combinations of responses were given a value of 2.

1	Yes	Heavy emphasis
2	No	No heavy emphasis
8	Omitted	Teacher omitted both questions

TREMP2 (Teacher heavy emphasis #2)

The responses to T046805 (How much instructional time is devoted to reading strategies?), T046902 (How much emphasis is given to the integration of reading and writing), and T046903 (How much emphasis is given to whole language?) were used to create TREMP2. The responses of "almost all of the time" for T046805 and "heavy emphasis" for T046902 and T046903 were considered. If the teacher provided the above responses to two or three of the questions, then TREMP2 was given a value of 1. Other combinations of responses were given a value of 2.

1	Yes	Heavy emphasis
2	No	No heavy emphasis
8	Omitted	Teacher omitted all three questions

TREMP3 (Teaching heavy emphasis #3)

The responses to T046904 (How much emphasis is given to reading across the content areas?) and T046905 (How much emphasis is given to individualized reading programs?) were used to create TREMP3. If the teacher responded "heavy emphasis" to T046904 and T046905, then TREMP3 was given a value of 1. Other combinations of responses were given a value of 2.

1	Yes	Heavy emphasis
2	No	No heavy emphasis
8	Omitted	Teacher omitted both questions

TRUMAJ (Teacher undergraduate major)

Items T040701 through T040705 in the teacher questionnaire (What were your undergraduate major fields of study?) were used to determine TRUMAJ as follows:

1	English/reading	The teacher responded yes to T040702 (English, reading, and/or language arts).
2	Education	The teacher responded yes to T040701 (education) and No to T040702.
3	Other	Any other response.

TRGMAJ (Teacher graduate major)

Items T040801 through T040806 in the teacher questionnaire (What were your undergraduate major fields of study?) were used to determine TRGMAJ as follows:

1	English/reading	The teacher responded yes to T040802 (English, reading, and/or language arts).
2	Education	The teacher responded yes to T040801 (education) and no to T040702.
3	Other	The teacher responded yes to T040803 (mathematics education), T040804 (mathematics), or T040805 (other).
4	None	The teacher indicated (T040806) that he or she had had no graduate-level study.

TRCERT (Type of teaching certification)

Items T040501 through T040505 (Do you have teaching certification in any of the following areas that is recognized by the state in which you teach?) were combined to create TRCERT. The following rules were used to determine the four values of TRCERT.

1	Reading	The teacher responded yes to T040502 (Reading)
2	Language arts	The teacher responded yes to T040503 (language arts) and no to T040502.
3	Education	The teacher responded yes to T040501 (education) and no to T040502 and T040503.
4	Other	Any other response

Variables Derived from the Mathematics Teacher Questionnaire

TCERTIF (Type of teaching certificate)

Items T040501 through T040505 (Do you have teaching certification in any of the following areas that is recognized by the state in which you teach?) in the teacher questionnaire were combined to produce TCERTIF. The following rules were used to determine the three values of TCERTIF.

1	Mathematics	The teacher responded Yes to T040504 (middle/junior high school or secondary mathematics)
2	Education	The teacher responded Yes to T040501 (elementary or middle/junior high school education [general]) and No to T040504
3	Other	Any other response

TUNDMJB (Undergraduate major)

Items T040701 through T040705 in the teacher questionnaire (What were your undergraduate major fields of study?) were used to determine TUNDMJB as follows:

1	Mathematics	The teacher responded Yes to T040703 (mathematics)
2	Mathematics education	The teacher responded Yes to T040704 (mathematics education) and No to T040703
3	Education	The teacher responded Yes to T040701 (education) and No to T040703 and T040704

4 Other	Any other response

TGRDMJB (Graduate major)

Items T040801 through T040806 in the teacher questionnaire (What were your graduate major fields of study?) were used to determine TGRDMJB as follows:

1 Mathematics	The teacher responded Yes to T040803 (mathematics)
2 Mathematics education	The teacher responded Yes to T040804 (mathematics education) and No to T040803
3 Education	The teacher responded Yes to T040801 (education) and No to T040803 and T040804
4 Other	Any other response

TMATHEX (Exposure in areas of mathematics)

Items T041602 through T041607 in the teacher questionnaire were used to determine teachers' exposure to the mathematics topics of number systems and numeration, measurement, geometry, probability/statistics, abstract/linear algebra, and calculus. For each area, teachers were asked to indicated up to four levels of exposure that applied to them: one or more college courses, part of a course, in-service training, and little or no exposure.

TMATHEX was derived by summing the areas for which teachers responded "one or more college courses," then categorizing the responses into three levels:

1 Five to six areas
2 Three to four areas
3 Zero to two areas

Variables Derived from the Writing Teacher Questionnaire

WTEIA (Teacher's emphasis on integrated approach)

Eighth-grade teachers' responses to T048803 (Instructional approach: integrating reading and writing), T048804 (Instructional approach: writing about literature), and T048805 (Instructional approach: writing across subjects) were combined to form WTEIA. The number of times a teacher responded to each of these questions with "Yes, as a central part of instruction" were summed. The range, therefore, is from 1 to 4, where 1 means none is central to the instructional approach, and 4 means all three are central to the instructional approach. If a response was omitted for any of the three items, the variable WTEIA was treated as omitted for that respondent.

WTEWP (Teacher's emphasis on the writing process)

Eighth-grade teachers' responses to T049201 (How often do you ask students to plan their writing?), T049203 (How often do you ask students to define purpose or audience?), T049208 (How often do you ask students to use other resources?), and T049206 (How often do you ask students to do more than one draft?) were combined to form WTEWP. The number of times a teacher responded to each of these questions with "Always" are summed. The range, therefore, is from 1 to 5, where 1 means not always asked to do any of the above writing processes, and 5 means all four writing processes are always asked of the students. If a response was omitted for any of the four items, the variable WTEWP was treated as omitted for that respondent.

WTENWP (Teacher's encouragement of the writing process)

Eighth-grade teachers' responses to T049201 (How often do you ask the students to plan writing?) and T049206 (How often do you ask the students to write more than 1 draft?) were combined to form WTENWP. The number of times a teacher responded "Always" to each of these questions were summed. The range, therefore, is 1 to 3, where 1 means that planning and writing more than one draft is not always asked, and 3 means that both planning and writing more than one draft are always done. If a response was omitted for either item, the variable WTENWP was treated as omitted for that respondent.

VARIABLES DERIVED FROM READING ITEMS

NORMIT (Normit Gaussian score)
SCHREAD (School-level mean Gaussian score)

The normit score is a student-level Gaussian score based on the inverse normal transformation of the mid-percentile rank of a student's number-correct booklet score within that booklet. The normit scores were used to decide collapsing of variables, finalize conditioning coding, and check the results of scaling.

The number correct is based on the number of dichotomous items answered correctly plus the score obtained on extended constructed-response items. The mid-percentile rank is based on the formula:

$$\frac{CF(i)+CF(i-1)}{2N}$$

where CF(i) is the cumulative frequency at i items correct and N is the total sample size. If i = 0 then

$$\frac{CF(0)+\frac{CF(1)}{2}}{2N}$$

A school-level normit, SCHREAD, was also created; this was the mean normit across all reading booklets administered in a school.

VARIABLES DERIVED FROM MATHEMATICS ITEMS

CALCUSE (Calculator-usage index)

In each calculator block, items were classified as calculator-suitable (items for which use of a calculator was either required or not inappropriate), and calculator-unsuitable (items for which use of calculator was inappropriate). For each item in a calculator block, students were asked to indicate whether or not they used a calculator in answering the items.

The 1992 examinees who were administered at least one of the calculator blocks were classified into two groups—"high" and "other." The "high" group consisted of those examinees who indicated that they had used a calculator for 65 percent or more of the calculator-suitable items that they attempted and had used a calculator on no more than one of the calculator-unsuitable items that they attempted. The "other" group consisted of everyone else. For the purpose of assigning students to the categories of this variable, interest was restricted to the set of items for which a student had indicated whether or not he or she had used a calculator. Items for which a student failed to indicate this were excluded from the calculation of percentages.

VARIABLES RELATED TO PROFICIENCY SCALING

Reading Proficiency Score Variables

Item response theory (IRT) was used to estimate average reading proficiency for the nation and for various subpopulations, based on students' performance on the set of reading items they received. IRT provides a common scale on which performance can be reported for the nation and for subpopulations, even when all students do not answer the same set of questions. This common scale makes it possible to report on relationships between students' characteristics (based on their responses to the background questions) and their overall performance in the assessment.

A scale ranging from 0 to 500 was created to report performance for each of the two content areas—Reading for Literary Experience and Reading to Gain Information. Each content-area scale was based on the distribution of student performance across all three grades assessed in the 1992 national assessment (grades 4, 8, and 12) and had a mean of 250 and a standard deviation of 50. A composite scale was created as an overall measure of students' mathematics proficiency. The composite scale was a weighted average of the content area scales, where the weight for each content area was proportional to the relative importance assigned to the content area as specified in the mathematics objectives.

Scale proficiency estimates were obtained for all students assessed in the 1992 reading assessments. The NAEP methods use random draws ("plausible values") from estimated proficiency distributions to compute population statistics. Plausible values are not optimal estimates of individual proficiency; instead, they serve as intermediate values to be used in estimating population characteristics. Chapter 11 provides further details on the computation and use of plausible values.

The reading proficiency score (plausible value) variables are named as shown in Table B-3.

Table B-3
Scaling Variables for the 1992 National Assessment Reading Samples

Reading Scale	Data Variables
Reading for Literary Experience	RRPS11 to RRPS15
Reading to Gain Information	RRPS21 to RRPS25
Reading to Perform a Task	RRPS31 to RRPS35
Composite	RRPCM1 to RRPCM5

Mathematics Proficiency Score Variables

Item response theory (IRT) was used to estimate average mathematics proficiency for the nation and for various subpopulations, based on students' performance on the set of mathematics items they received. IRT provides a common scale on which performance can be reported for the nation and subpopulations, even when all students do not answer the same set of questions. This common scale makes it possible to report on relationships between students' characteristics (based on their responses to the background questions) and their overall performance in the assessment.

A scale ranging from 0 to 500 was created to report performance for each of the five mathematics content areas: Numbers and Operations; Measurement, Geometry; Data Analysis, Statistics, and Probability; and Algebra and Functions. A composite scale was created as an overall measure of students' mathematics proficiency. The composite scale was a weighted average of the five content area scales, where the weight for each content area was proportional to the relative importance assigned to the content area as specified in the mathematics objectives. An additional scale was created for the items designed to measure estimation abilities.

Scale proficiency estimates were obtained for all students assessed in 1992 mathematics assessments. The NAEP methods use random draws ("plausible values") from estimated proficiency distributions to compute population statistics. Plausible values are not optimal estimates of individual proficiency; instead, they serve as intermediate values to be used in estimating population characteristics. Chapter 11 provides further details on the computation and use of plausible values.

The mathematics proficiency score (plausible value) variables are named as shown in Table B-4.

Table B-4
Scaling Variables for the 1992 National Assessment Mathematics Samples

Mathematics Scale	Data Variables
Numbers and Operations	MRPS11 to MRPS15
Measurement	MRPS21 to MRPS25
Geometry	MRPS31 to MRPS35
Data Analysis, Statistics, and Probability	MRPS41 to MRPS45
Algebra and Functions	MRPS51 to MRPS55
Composite	MRPCM1 to MRPCM5
Estimation	MRPES1 to MRPES5

PRINCIPAL'S QUESTIONNAIRE VARIABLES (PQ)

Before the assessment, Westat, Inc., distributed a questionnaire to the principal of each participating school to gather data about school characteristics, including parents' occupations and student race/ethnicity. The data variables from this questionnaire are retained on the school file. A subset of these variables are also on the student files. Principal's questionnaire variables are identified in the data layouts by "(PQ)" in the SHORT LABEL field.

QUALITY EDUCATION DATA VARIABLES (QED)

The data files contain several variables obtained from information supplied by Quality Education Data, Inc. (QED). QED maintains and updates annually lists of schools showing grade span, total enrollment, instructional dollars per pupil, and other information for each school. These data variables are retained on both the school and student files and are identified in the data layouts by "(QED)" in the SHORT LABEL field.

Most of the QED variables are defined sufficiently in the data codebooks. Explanations of others are provided below.

ORSHPT and SORSHPT are the Orshansky Percentile, an indicator of relative wealth that specifies the percentage of school-age children in a district who fall below the poverty line.

IDP and SIDP represent, at the school district level, dollars per student spent for textbooks and supplemental materials.

ADULTED and SADLTED indicate whether or not adult education courses are offered at the school site.

URBAN and SURBAN define the school's urbanicity: urban (central city); suburban (area surrounding central city, but still located within the counties constituting the metropolitan statistical area); or rural (area outside any metropolitan statistical area).

APPENDIX C

Distribution of Weight Components for 1992 NAEP Samples

Appendix C

DISTRIBUTION OF WEIGHT COMPONENTS FOR 1992 NAEP SAMPLES

The following tables, which are cited and described throughout Chapter 10, "Weighting Procedures and Estimation of Sampling Variance," show the distribution of student and excluded student weight components for the 1992 NAEP samples, including base weights, the various nonresponse adjustment factors, trimming factors, and poststratification factors.

Table C-1
Distribution of Student Base Weights, Long-term Trend Samples

Sample	Number of Cases	Mean	Standard Deviation	Minimum	25th Percentile	Median	75th Percentile	Maximum
AGE 9 LONG-TERM TREND SAMPLES								
Booklets 51-56	7062	612.6	2.6	250.2	420.7	620.3	739.3	2604.9
Booklets 91-93	7335	408.7	1.2	193.9	358.2	395.1	442.6	1738.8
Excluded students	935	336.1	6.6	118.5	230.1	256.4	329.0	1311.5
AGE 13 LONG-TERM TREND SAMPLES								
Booklets 51-56	5514	726.5	5.3	310.2	461.0	672.8	804.0	2921.5
Booklets 91-93	5909	463.7	2.2	190.8	373.3	429.5	468.4	1989.8
Excluded students	708	391.5	12.3	114.5	237.3	280.5	359.3	2905.4
AGE 17 LONG-TERM TREND SAMPLES								
Booklets 51-56	5569	485.0	3.6	58.4	305.0	394.2	608.6	1804.8
Booklets 84-85	4359	483.8	4.7	117.2	299.8	387.2	594.8	2936.2
Excluded students	623	288.4	7.0	107.9	193.3	230.6	301.0	1481.1

Table C-2
Distribution of Student Base Weights, Main Samples

Sample	Number of Cases	Mean	Standard Deviation	Minimum	25th Percentile	Median	75th Percentile	Maximum
AGE 9/GRADE 4 MAIN SAMPLES								
Booklets 1-26	9414	443.3	2.5	136.9	249.7	365.1	608.4	1300.2
Booklet 27	2054	1937.4	18.7	823.4	1369.0	1625.8	2737.9	3998.8
Booklet 28	2236	2000.3	19.1	742.8	1349.2	1693.7	2737.9	3921.8
Booklets 30-45	8416	497.7	3.2	121.1	260.8	400.3	685.1	2819.7
Booklets 60-77	9552	425.7	2.4	105.3	228.2	361.7	554.3	1372.7
Excluded students	2990	123.5	1.0	48.0	94.5	109.0	133.6	444.7
AGE 13/GRADE 8 MAIN SAMPLES								
Booklets 1-26	10291	376.7	2.4	93.9	200.3	284.2	467.6	2230.7
Booklet 27	2416	1771.1	19.9	703.9	1089.9	1513.0	2111.8	8447.2
Booklets 30-49	14942	257.7	1.3	61.9	143.7	209.6	336.7	1871.2
Booklets 60-79	14942	263.8	1.3	67.3	148.0	209.3	351.9	1598.8
Excluded students	3400	89.2	0.9	38.1	64.1	70.2	115.3	561.4
AGE 17/GRADE 12 MAIN SAMPLES								
Booklets 1-26	9499	280.3	2.0	53.5	132.8	224.9	353.4	1449.0
Booklet 27	2074	1134.0	13.6	334.1	677.4	921.9	1536.6	3502.4
Booklets 30-50	15314	172.0	0.9	37.5	97.8	140.2	226.3	743.5
Booklets 60-80	15669	166.1	0.9	40.9	87.0	128.2	231.1	781.1
Excluded students	2926	62.8	0.6	15.9	37.3	49.1	82.2	274.1

Table C-3
Distribution of School Nonresponse Adjustments, Long-term Trend Samples

Sample	Number of Cases	Mean	Standard Deviation	Minimum	25th Percentile	Median	75th Percentile	Maximum
AGE 9 LONG-TERM TREND SAMPLES								
Booklets 51-56	7062	1.0	0.0	1.0	1.0	1.0	1.0	1.0
Booklets 91-93	7335	1.0	0.0	1.0	1.0	1.0	1.0	1.0
Excluded students	935	1.0	0.0	1.0	1.0	1.0	1.0	1.0
AGE 13 LONG-TERM TREND SAMPLES								
Booklets 51-56	5514	1.0	0.0	1.0	1.0	1.0	1.0	1.0
Booklets 91-93	5909	1.0	0.0	1.0	1.0	1.0	1.0	1.0
Excluded students	708	1.0	0.0	1.0	1.0	1.0	1.0	1.0
AGE 17 LONG-TERM TREND SAMPLES								
Booklets 51-56	5569	1.2	0.0	1.0	1.1	1.2	1.3	1.4
Booklets 84-85	4359	1.2	0.0	1.0	1.1	1.2	1.3	1.4
Excluded students	623	1.2	0.0	1.0	1.1	1.2	1.3	1.4

Table C-4
Distribution of School Nonresponse Adjustments, Main Samples

Sample	Number of Cases	Mean	Standard Deviation	Minimum	25th Percentile	Median	75th Percentile	Maximum
AGE 9/GRADE 4 MAIN SAMPLES								
Booklets 1-26	9414	1.0	0.0	1.0	1.0	1.0	1.0	1.0
Booklet 27	2054	1.0	0.0	1.0	1.0	1.0	1.0	1.0
Booklet 28	2236	1.0	0.0	1.0	1.0	1.0	1.0	1.0
Booklets 30-45	8416	1.0	0.0	1.0	1.0	1.0	1.0	1.0
Booklets 60-77	9552	1.0	0.0	1.0	1.0	1.0	1.0	1.0
Excluded students	2990	1.0	0.0	1.0	1.0	1.0	1.0	1.0
AGE 13/GRADE 8 MAIN SAMPLES								
Booklets 1-26	10291	1.0	0.0	1.0	1.0	1.0	1.0	1.1
Booklet 27	2416	1.0	0.0	1.0	1.0	1.0	1.0	1.1
Booklets 30-49	14942	1.0	0.0	1.0	1.0	1.0	1.0	1.1
Booklets 60-79	14942	1.0	0.0	1.0	1.0	1.0	1.0	1.1
Excluded students	3400	1.0	0.0	1.0	1.0	1.0	1.0	1.1
AGE 17/GRADE 12 MAIN SAMPLES								
Booklets 1-26	9499	1.2	0.0	1.0	1.1	1.2	1.3	1.3
Booklet 27	2074	1.2	0.0	1.0	1.1	1.2	1.3	1.3
Booklets 30-50	15314	1.2	0.0	1.0	1.1	1.2	1.3	1.3
Booklets 60-80	15669	1.2	0.0	1.0	1.1	1.2	1.2	1.3
Excluded students	2926	1.2	0.0	1.0	1.1	1.2	1.3	1.3

Table C-5
Distribution of Session Nonresponse Adjustments, Long-term Trend Samples

Sample	Number of Cases	Mean	Standard Deviation	Minimum	25th Percentile	Median	75th Percentile	Maximum
AGE 9 LONG-TERM TREND SAMPLES								
Booklets 51-56	7062	1.1	0.0	1.0	1.1	1.2	1.2	1.3
Booklets 91-93	7335	1.1	0.0	1.0	1.0	1.1	1.2	1.3
Excluded students	935	1.2	0.0	1.0	1.1	1.2	1.2	1.3
AGE 13 LONG-TERM TREND SAMPLES								
Booklets 51-56	5514	1.2	0.0	1.0	1.0	1.2	1.3	1.4
Booklets 91-93	5909	1.2	0.0	1.0	1.0	1.2	1.3	1.4
Excluded students	708	1.2	0.0	1.0	1.0	1.2	1.3	1.4
AGE 17 LONG-TERM TREND SAMPLES								
Booklets 51-56	5569	1.2	0.0	1.1	1.1	1.2	1.2	1.6
Booklets 84-85	4359	1.2	0.0	1.1	1.1	1.2	1.2	1.6
Excluded students	623	1.2	0.0	1.1	1.1	1.2	1.2	1.6

Table C-6
Distribution of Session Nonresponse Adjustments, Main Samples

Sample	Number of Cases	Mean	Standard Deviation	Minimum	25th Percentile	Median	75th Percentile	Maximum
AGE 9/GRADE 4 MAIN SAMPLES								
Booklets 1-26	9414	1.2	0.0	1.0	1.0	1.1	1.2	1.9
Booklet 27	2054	1.1	0.0	1.0	1.0	1.2	1.2	1.6
Booklet 28	2236	1.1	0.0	1.0	1.0	1.1	1.2	1.4
Booklets 30-45	8416	1.1	0.0	1.0	1.0	1.1	1.2	1.5
Booklets 60-77	9552	1.2	0.0	1.0	1.0	1.1	1.2	1.6
Excluded students	2990	1.2	0.0	1.0	1.0	1.2	1.2	1.7
AGE 13/GRADE 8 MAIN SAMPLES								
Booklets 1-26	10291	1.1	0.0	1.0	1.1	1.1	1.2	1.8
Booklet 27	2416	1.2	0.0	1.0	1.0	1.2	1.3	1.5
Booklets 30-49	14942	1.2	0.0	1.0	1.1	1.2	1.2	2.0
Booklets 60-79	14942	1.2	0.0	1.0	1.1	1.2	1.2	1.7
Excluded students	3400	1.1	0.0	1.0	1.1	1.2	1.2	1.8
AGE 17/GRADE 12 MAIN SAMPLES								
Booklets 1-26	9499	1.2	0.0	1.0	1.1	1.1	1.3	1.5
Booklet 27	2074	1.1	0.0	1.0	1.0	1.1	1.2	1.3
Booklets 30-50	15314	1.2	0.0	1.0	1.1	1.2	1.3	1.6
Booklets 60-80	15669	1.2	0.0	1.0	1.0	1.2	1.3	1.5
Excluded students	2926	1.2	0.0	1.0	1.1	1.2	1.3	1.5

Table C-7
Distribution of Student Nonresponse Adjustments, Long-term Trend Samples

Sample	Number of Cases	Mean	Standard Deviation	Minimum	25th Percentile	Median	75th Percentile	Maximum
AGE 9 LONG-TERM TREND SAMPLES								
Booklets 51-56	7062	1.1	0.0	1.0	1.1	1.1	1.1	1.1
Booklets 91-93	7335	1.1	0.0	1.0	1.0	1.1	1.1	1.2
Excluded students	935	1.0	0.0	1.0	1.0	1.0	1.0	1.4
AGE 13 LONG-TERM TREND SAMPLES								
Booklets 51-56	5514	1.1	0.0	1.0	1.1	1.1	1.1	1.3
Booklets 91-93	5909	1.1	0.0	1.1	1.1	1.1	1.1	1.3
Excluded students	708	1.0	0.0	1.0	1.0	1.0	1.0	1.0
AGE 17 LONG-TERM TREND SAMPLES								
Booklets 51-56	5569	1.2	0.0	1.1	1.1	1.2	1.2	1.4
Booklets 84-85	4359	1.2	0.0	1.1	1.2	1.2	1.2	1.7
Excluded students	623	1.1	0.0	1.0	1.0	1.1	1.2	1.3

Table C-8
Distribution of Student Nonresponse Adjustments, Main Samples

Sample	Number of Cases	Mean	Standard Deviation	Minimum	25th Percentile	Median	75th Percentile	Maximum
AGE 9/GRADE 4 MAIN SAMPLES								
Booklets 1-26	9414	1.1	0.0	1.0	1.0	1.1	1.1	1.4
Booklet 27	2054	1.1	0.0	1.0	1.0	1.1	1.1	1.2
Booklet 28	2236	1.1	0.0	1.0	1.0	1.0	1.1	1.2
Booklets 30-45	8416	1.1	0.0	1.0	1.0	1.1	1.1	1.2
Booklets 60-77	9552	1.1	0.0	1.0	1.0	1.1	1.1	1.3
Excluded students	2990	1.1	0.0	1.0	1.0	1.1	1.2	1.3
AGE 13/GRADE 8 MAIN SAMPLES								
Booklets 1-26	10291	1.1	0.0	1.0	1.1	1.1	1.1	1.5
Booklet 27	2416	1.1	0.0	1.0	1.1	1.1	1.1	1.5
Booklets 30-49	14942	1.1	0.0	1.0	1.1	1.1	1.1	1.3
Booklets 60-79	14942	1.1	0.0	1.0	1.1	1.1	1.2	1.4
Excluded students	3400	1.1	0.0	1.0	1.0	1.1	1.2	1.3
AGE 17/GRADE 12 MAIN SAMPLES								
Booklets 1-26	9499	1.2	0.0	1.1	1.2	1.2	1.3	1.7
Booklet 27	2074	1.3	0.0	1.0	1.2	1.2	1.3	1.7
Booklets 30-50	15314	1.2	0.0	1.0	1.2	1.2	1.3	1.5
Booklets 60-80	15669	1.2	0.0	1.1	1.2	1.2	1.3	1.5
Excluded students	2926	1.1	0.0	1.0	1.0	1.0	1.2	1.3

Table C-9
Distribution of Trimming Factors, Long-term Trend Samples

Sample	Number of Cases	Mean	Standard Deviation	Minimum	25th Percentile	Median	75th Percentile	Maximum
AGE 9 LONG-TERM TREND SAMPLES								
Booklets 51-56	7062	1.0	0.0	0.7	1.0	1.0	1.0	1.0
Booklets 91-93	7335	1.0	0.0	0.8	1.0	1.0	1.0	1.0
Excluded students	935	1.0	0.0	0.6	1.0	1.0	1.0	1.0
AGE 13 LONG-TERM TREND SAMPLES								
Booklets 51-56	5514	1.0	0.0	0.5	1.0	1.0	1.0	1.0
Booklets 91-93	5909	1.0	0.0	0.6	1.0	1.0	1.0	1.0
Excluded students	708	1.0	0.0	0.8	1.0	1.0	1.0	1.0
AGE 17 LONG-TERM TREND SAMPLES								
Booklets 51-56	5569	1.0	0.0	0.5	1.0	1.0	1.0	1.0
Booklets 84-85	4359	1.0	0.0	0.5	1.0	1.0	1.0	1.0
Excluded students	623	1.0	0.0	0.6	1.0	1.0	1.0	1.0

Table C-10
Distribution of Trimming Factors, Main Samples

Sample	Number of Cases	Mean	Standard Deviation	Minimum	25th Percentile	Median	75th Percentile	Maximum
AGE 9/GRADE 4 MAIN SAMPLES								
Booklets 1-26	9414	1.0	0.0	0.9	1.0	1.0	1.0	1.0
Booklet 27	2054	1.0	0.0	1.0	1.0	1.0	1.0	1.0
Booklet 28	2236	1.0	0.0	0.9	1.0	1.0	1.0	1.0
Booklets 30-45	8416	1.0	0.0	0.2	1.0	1.0	1.0	1.0
Booklets 60-77	9552	1.0	0.0	0.5	1.0	1.0	1.0	1.0
Excluded students	2990	1.0	0.0	0.5	1.0	1.0	1.0	1.0
AGE 13/GRADE 8 MAIN SAMPLES								
Booklets 1-26	10291	1.0	0.0	0.5	1.0	1.0	1.0	1.0
Booklet 27	2416	1.0	0.0	0.7	1.0	1.0	1.0	1.0
Booklets 30-49	14942	1.0	0.0	0.5	1.0	1.0	1.0	1.0
Booklets 60-79	14942	1.0	0.0	0.4	1.0	1.0	1.0	1.0
Excluded students	3400	1.0	0.0	0.7	1.0	1.0	1.0	1.0
AGE 17/GRADE 12 MAIN SAMPLES								
Booklets 1-26	9499	1.0	0.0	0.6	1.0	1.0	1.0	1.0
Booklet 27	2074	1.0	0.0	0.7	1.0	1.0	1.0	1.0
Booklets 30-50	15314	1.0	0.0	0.6	1.0	1.0	1.0	1.0
Booklets 60-80	15669	1.0	0.0	0.6	1.0	1.0	1.0	1.0
Excluded students	2926	0.9	0.0	0.4	1.0	1.0	1.0	1.0

Table C-11
Distribution of Poststratification Factors, Long-term Trend Samples

Sample	Number of Cases	Mean	Standard Deviation	Minimum	25th Percentile	Median	75th Percentile	Maximum
AGE 9 LONG-TERM TREND SAMPLES								
Booklets 51-56	7062	1.1	0.0	0.7	1.0	1.0	1.2	1.6
Booklets 91-93	7335	1.1	0.0	0.7	1.1	1.1	1.2	1.2
Excluded students	935	1.1	0.0	0.7	0.8	1.0	1.3	1.6
AGE 13 LONG-TERM TREND SAMPLES								
Booklets 51-56	5514	1.1	0.0	0.7	0.9	1.0	1.3	1.5
Booklets 91-93	5909	1.1	0.0	0.7	1.0	1.1	1.1	1.1
Excluded students	708	1.1	0.0	0.8	0.9	1.2	1.3	1.5
AGE 17 LONG-TERM TREND SAMPLES								
Booklets 51-56	5569	1.3	0.0	0.9	1.1	1.2	1.4	2.3
Booklets 84-85	4359	1.2	0.0	0.8	0.9	1.2	1.3	1.5
Excluded students	623	1.4	0.0	0.9	1.1	1.4	1.7	2.3

Table C-12
Distribution of Poststratification Factors, Main Samples

Sample	Number of Cases	Mean	Standard Deviation	Minimum	25th Percentile	Median	75th Percentile	Maximum
AGE 9/GRADE 4 MAIN SAMPLES								
Booklets 1-26	9414	1.1	0.0	0.6	1.0	1.1	1.3	1.7
Booklet 27	2054	1.2	0.0	0.7	0.8	1.1	1.5	2.5
Booklet 28	2236	1.0	0.0	0.7	0.9	1.0	1.1	1.7
Booklets 30-45	8416	1.1	0.0	0.6	1.0	1.1	1.3	1.8
Booklets 60-77	9552	1.1	0.0	0.7	1.0	1.2	1.3	1.8
Excluded students	2990	1.0	0.0	0.6	0.8	1.0	1.3	1.7
AGE 13/GRADE 8 MAIN SAMPLES								
Booklets 1-26	10291	1.1	0.0	0.8	1.0	1.0	1.2	1.6
Booklet 27	2416	1.0	0.0	0.6	0.7	1.0	1.2	1.9
Booklets 30-49	14942	1.1	0.0	0.8	0.9	1.0	1.2	1.6
Booklets 60-79	14942	1.1	0.0	0.7	0.9	1.0	1.2	1.5
Excluded students	3400	1.1	0.0	0.8	0.9	1.0	1.2	1.6
AGE 17/GRADE 12 MAIN SAMPLES								
Booklets 1-26	9499	1.1	0.0	0.9	1.0	1.1	1.2	1.6
Booklet 27	2074	1.3	0.0	0.9	1.1	1.3	1.4	1.6
Booklets 30-50	15314	1.1	0.0	0.9	1.0	1.1	1.1	1.6
Booklets 60-80	15669	1.2	0.0	0.9	1.1	1.1	1.2	1.5
Excluded students	2926	1.2	0.0	0.9	1.1	1.1	1.4	1.5

Table C-13
Distribution of Aggregate Adjustments to Base Weights, Long-term Trend Samples

Sample	Number of Cases	Mean	Standard Deviation	Minimum	25th Percentile	Median	75th Percentile	Maximum
AGE 9 LONG-TERM TREND SAMPLES								
Booklets 51-56	7062	1.3	0.0	0.7	1.1	1.3	1.5	2.5
Booklets 91-93	7335	1.3	0.0	0.7	1.2	1.3	1.4	1.7
Excluded students	935	1.3	0.0	0.5	1.0	1.2	1.5	2.3
AGE 13 LONG-TERM TREND SAMPLES								
Booklets 51-56	5514	1.4	0.0	0.5	1.1	1.4	1.7	2.4
Booklets 91-93	5909	1.4	0.0	0.5	1.2	1.3	1.5	2.0
Excluded students	708	1.4	0.0	0.7	1.2	1.4	1.6	2.1
AGE 17 LONG-TERM TREND SAMPLES								
Booklets 51-56	5569	2.1	0.0	0.6	1.7	2.0	2.5	4.9
Booklets 84-85	4359	2.0	0.0	0.8	1.6	1.9	2.3	4.3
Excluded students	623	2.2	0.0	0.9	1.7	2.0	2.7	4.6

Table C-14
Distribution of Aggregate Adjustments to Base Weights, Main Samples

Sample	Number of Cases	Mean	Standard Deviation	Minimum	25th Percentile	Median	75th Percentile	Maximum
AGE 9/GRADE 4 MAIN SAMPLES								
Booklets 1-26	9414	1.4	0.0	0.6	1.1	1.3	1.6	2.9
Booklet 27	2054	1.4	0.0	0.7	1.1	1.3	1.7	3.2
Booklet 28	2236	1.3	0.0	0.7	1.0	1.2	1.5	2.9
Booklets 30-45	8416	1.4	0.0	0.3	1.1	1.3	1.5	2.6
Booklets 60-77	9552	1.4	0.0	0.5	1.2	1.4	1.6	3.3
Excluded students	2990	1.3	0.0	0.5	0.9	1.2	1.6	3.1
AGE 13/GRADE 8 MAIN SAMPLES								
Booklets 1-26	10291	1.4	0.0	0.4	1.1	1.3	1.6	3.0
Booklet 27	2416	1.3	0.0	0.5	1.0	1.2	1.6	2.5
Booklets 30-49	14942	1.4	0.0	0.4	1.1	1.3	1.6	3.4
Booklets 60-79	14942	1.4	0.0	0.4	1.1	1.3	1.6	2.5
Excluded students	3400	1.4	0.0	0.7	1.2	1.3	1.6	2.4
AGE 17/GRADE 12 MAIN SAMPLES								
Booklets 1-26	9499	2.0	0.0	1.1	1.6	1.9	2.2	5.3
Booklet 27	2074	2.1	0.0	1.0	1.6	2.1	2.4	3.6
Booklets 30-50	15314	2.0	0.0	1.0	1.7	1.9	2.3	4.6
Booklets 60-80	15669	2.0	0.0	1.0	1.7	2.0	2.3	3.9
Excluded students	2926	1.8	0.0	0.6	1.4	1.7	2.1	3.3

Table C-15
Distribution of Final Student Weights, Long-term Trend Samples

Sample	Number of Cases	Mean	Standard Deviation	Minimum	25th Percentile	Median	75th Percentile	Maximum
AGE 9 LONG-TERM TREND SAMPLES								
Booklets 51-56	7062	707.9	3.3	183.7	487.7	693.7	867.8	3317.6
Booklets 91-93	7335	458.4	1.7	135.6	379.4	447.4	533.2	2167.7
Excluded students	935	368.3	7.9	93.3	215.9	289.7	405.4	1686.1
AGE 13 LONG-TERM TREND SAMPLES								
Booklets 51-56	5514	854.7	6.0	253.2	580.8	756.5	1041.2	4187.4
Booklets 91-93	5909	538.4	2.4	202.0	426.0	505.6	582.5	2295.5
Excluded students	708	449.2	13.7	113.6	269.9	323.6	439.1	3585.7
AGE 17 LONG-TERM TREND SAMPLES								
Booklets 51-56	5569	729.8	5.6	60.9	420.1	621.1	898.6	3585.3
Booklets 84-85	4359	673.7	6.0	169.5	417.2	547.7	822.9	4192.4
Excluded students	623	449.5	12.0	112.3	257.6	358.9	494.3	2030.1

Table C-16
Distribution of Final Student Weights, Main Samples

Sample	Number of Cases	Mean	Standard Deviation	Minimum	25th Percentile	Median	75th Percentile	Maximum
AGE 9/GRADE 4 MAIN SAMPLES								
Booklets 1-26	9414	522.9	3.3	90.4	288.1	414.7	712.4	2097.1
Booklet 27	2054	2379.6	28.5	572.5	1418.1	2042.8	3031.4	8722.2
Booklet 28	2236	2205.7	23.3	620.9	1327.9	1881.5	2984.0	7037.8
Booklets 30-45	8416	583.7	3.8	90.7	312.7	477.1	798.3	3749.7
Booklets 60-77	9552	513.9	3.0	74.9	288.4	426.5	674.8	1846.4
Excluded students	2990	142.0	1.6	29.9	85.9	115.9	175.9	852.3
AGE 13/GRADE 8 MAIN SAMPLES								
Booklets 1-26	10291	451.8	3.0	85.4	238.3	363.2	587.8	3168.2
Booklet 27	2416	1940.4	21.9	521.8	1180.8	1635.8	2401.7	9203.8
Booklets 30-49	14942	311.4	1.6	54.6	180.1	258.6	392.1	3254.6
Booklets 60-79	14942	312.1	1.6	56.8	176.3	255.4	402.5	3011.2
Excluded students	3400	107.0	1.1	31.0	66.5	87.8	128.9	767.1
AGE 17/GRADE 12 MAIN SAMPLES								
Booklets 1-26	9499	378.6	2.4	59.7	195.1	319.1	493.0	1982.0
Booklet 27	2074	1740.7	20.2	385.8	1019.0	1456.7	2273.5	5137.5
Booklets 30-50	15314	235.5	1.0	42.1	140.3	203.8	308.1	979.2
Booklets 60-80	15669	230.7	1.0	46.2	129.1	197.1	312.8	932.8
Excluded students	2926	75.4	0.7	15.5	47.9	65.5	94.1	423.0

696

695

APPENDIX D

Design Effects Statistics for 1992 NAEP Samples

Table D-1

Distributions of Design Effects by Demographic Subgroups for Proportion-correct Statistics Across Cognitive Reading Items in the Main Samples

Grade 4*

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.26	1.50	1.73	2.43	1.50	0.37	33
Male	1.22	1.40	1.60	2.26	1.42	0.32	39
Female	1.16	1.40	1.65	2.14	1.41	0.30	42
White	1.21	1.43	1.70	2.59	1.46	0.36	32
Black	1.04	1.25	1.44	2.08	1.28	0.31	33
Hispanic	1.16	1.34	1.55	2.09	1.37	0.29	44
Asian American	0.86	1.10	1.30	1.97	1.11	0.32	23
Other Race/ethnicity	1.06	1.25	1.47	2.23	1.28	0.31	34
Other Metro	1.14	1.48	1.73	2.45	1.46	0.39	27
Disadvantaged Urban	1.18	1.35	1.69	2.53	1.44	0.40	26
Advantaged Urban	1.09	1.50	1.95	4.56	1.60	0.76	9
Par. Ed. < HS	1.18	1.31	1.51	2.09	1.35	0.28	47
Par. Ed. = HS	1.11	1.32	1.56	1.94	1.32	0.31	35
Par. Ed. > HS	1.02	1.23	1.52	2.48	1.32	0.42	19
Par. Ed. = Coll.	1.22	1.40	1.66	2.65	1.44	0.37	31
Par. Ed. = IDK	1.09	1.29	1.62	2.65	1.36	0.38	26
Public School	1.23	1.43	1.67	2.20	1.44	0.36	32
Nonpublic School	1.04	1.27	1.59	3.33	1.36	0.46	18

---

* Distributions are based on 85 items.

Table D-2

Distributions of Design Effects by Demographic Subgroups for Proportion-correct Statistics Across Cognitive Reading Items in the Main Samples

Grade 8*

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.14	1.31	1.59	2.59	1.38	0.35	31
Male	1.04	1.29	1.54	2.47	1.31	0.33	31
Female	1.03	1.22	1.53	2.72	1.31	0.37	25
White	1.09	1.31	1.53	3.11	1.36	0.37	26
Black	0.93	1.14	1.44	2.50	1.20	0.34	25
Hispanic	0.94	1.12	1.34	3.20	1.18	0.36	21
Asian American	1.04	1.27	1.58	2.82	1.34	0.45	18
Other Race/ethnicity	0.95	1.16	1.44	2.52	1.22	0.38	20
Other Metro	1.09	1.34	1.61	2.86	1.40	0.38	27
Disadvantaged Urban	0.89	1.19	1.48	2.55	1.23	0.43	16
Advantaged Urban	0.87	1.19	1.58	2.61	1.27	0.49	13
Par. Ed. < HS	1.06	1.24	1.48	2.42	1.28	0.35	26
Par. Ed. = HS	0.99	1.17	1.43	2.53	1.21	0.34	26
Par. Ed. > HS	1.00	1.24	1.44	2.16	1.26	0.34	28
Par. Ed. = Coll.	1.06	1.26	1.50	2.15	1.30	0.32	33
Par. Ed. = IDK	1.05	1.25	1.42	2.12	1.25	0.29	38
Public School	1.10	1.26	1.49	2.46	1.32	0.33	31
Nonpublic School	1.12	1.33	1.64	2.83	1.41	0.41	23

---

* Distributions are based on 123 items.

Table D-3

Distributions of Design Effects by Demographic Subgroups for Proportion-correct Statistics Across Cognitive Reading Items in the Main Samples

Grade 12*

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.13	1.30	1.51	2.69	1.35	0.34	31
Male	1.09	1.23	1.51	2.08	1.30	0.30	37
Female	1.06	1.26	1.46	2.32	1.27	0.30	35
White	1.10	1.30	1.47	2.89	1.33	0.35	29
Black	1.04	1.25	1.52	2.47	1.29	0.34	28
Hispanic	1.07	1.33	1.75	3.43	1.47	0.58	13
Asian American	1.04	1.30	1.48	2.95	1.31	0.40	21
Other Race/ethnicity	0.90	1.17	1.43	3.72	1.21	0.46	14
Other Metro	1.08	1.29	1.53	2.90	1.34	0.36	27
Disadvantaged Urban	1.11	1.34	1.62	2.34	1.39	0.39	25
Advantaged Urban	0.98	1.29	1.60	2.93	1.32	0.44	17
Par. Ed. < HS	1.05	1.23	1.38	2.15	1.24	0.30	34
Par. Ed. = HS	0.98	1.13	1.30	1.87	1.16	0.26	40
Par. Ed. > HS	1.04	1.19	1.42	2.17	1.25	0.30	34
Par. Ed. = Coll.	1.11	1.33	1.53	2.37	1.34	0.29	41
Par. Ed. = IDK	1.01	1.22	1.44	2.10	1.26	0.31	32
Public School	1.09	1.29	1.46	2.46	1.31	0.32	34
Nonpublic School	1.05	1.19	1.50	2.28	1.28	0.36	26

---

* Distributions are based on 133 items.

Table D-4

Distributions of Design Effects by Demographic Subgroups for Proportion-correct Statistics Across Cognitive Reading Items in the Long-term Trend Samples

Age 9*

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	0.99	1.14	1.37	2.05	1.18	0.29	33
Male	0.91	1.14	1.29	2.13	1.14	0.29	30
Female	0.96	1.14	1.31	2.64	1.19	0.34	25
White	0.98	1.20	1.37	2.01	1.19	0.31	30
Black	0.88	1.11	1.35	2.05	1.15	0.34	22
Hispanic	0.90	1.09	1.32	2.41	1.14	0.34	22
Asian American	0.78	1.06	1.43	3.75	1.21	0.65	7
Other Race/ethnicity	1.02	1.18	1.35	2.35	1.19	0.31	30
Other Metro	0.95	1.13	1.37	2.05	1.18	0.31	29
Disadvantaged Urban	0.74	0.97	1.32	2.48	1.06	0.42	12
Advantaged Urban	0.90	1.12	1.44	2.73	1.21	0.43	16
Par. Ed. < HS	0.89	1.07	1.25	1.91	1.08	0.27	31
Par. Ed. = HS	0.91	1.08	1.38	2.00	1.15	0.32	25
Par. Ed. > HS	0.93	1.03	1.29	1.85	1.10	0.27	33
Par. Ed. = Coll.	1.02	1.15	1.31	2.08	1.18	0.27	37
Par. Ed. = IDK	0.98	1.13	1.32	1.80	1.15	0.26	38
Public School	0.94	1.17	1.36	2.08	1.17	0.28	34
Nonpublic School	0.92	1.14	1.44	2.46	1.22	0.42	16

---

* Distributions are based on 105 items.

Table D-5

Distributions of Design Effects by Demographic Subgroups for Proportion-correct Statistics Across Cognitive Reading Items in the Long-term Trend Samples

Age 13*

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.06	1.27	1.49	2.63	1.32	0.37	25
Male	1.02	1.24	1.50	2.35	1.28	0.36	25
Female	1.06	1.23	1.54	2.13	1.29	0.33	29
White	1.02	1.20	1.50	2.89	1.31	0.41	20
Black	1.02	1.28	1.54	2.46	1.31	0.41	20
Hispanic	0.84	1.06	1.29	2.80	1.12	0.41	15
Asian American	0.76	0.99	1.25	4.53	1.17	0.72	5
Other Race/ethnicity	0.74	0.98	1.21	2.21	1.01	0.40	12
Other Metro	1.04	1.26	1.51	2.69	1.30	0.40	21
Disadvantaged Urban	1.03	1.30	1.67	2.81	1.35	0.48	16
Advantaged Urban	0.87	1.28	1.81	5.56	1.49	0.92	5
Par. Ed. < HS	0.90	1.06	1.31	2.02	1.11	0.30	27
Par. Ed. = HS	1.04	1.23	1.47	2.75	1.27	0.35	26
Par. Ed. > HS	1.01	1.17	1.41	2.08	1.20	0.31	30
Par. Ed. = Coll.	1.11	1.31	1.54	2.67	1.34	0.34	31
Par. Ed. = IDK	0.98	1.21	1.50	2.69	1.25	0.35	25
Public School	1.08	1.23	1.57	2.78	1.33	0.38	24
Nonpublic School	1.02	1.16	1.49	3.09	1.31	0.47	15

---

* Distributions are based on 107 items.

Table D-6

Distributions of Design Effects by Demographic Subgroups for Proportion-correct Statistics Across Cognitive Reading Items in the Long-term Trend Samples

Age 17*

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.07	1.26	1.49	2.41	1.29	0.33	31
Male	1.09	1.28	1.53	2.18	1.31	0.31	35
Female	1.00	1.17	1.38	2.04	1.21	0.31	30
White	1.05	1.30	1.60	2.25	1.35	0.37	26
Black	0.93	1.11	1.42	1.91	1.14	0.29	30
Hispanic	0.75	0.93	1.17	1.84	0.98	0.32	19
Asian American	0.72	0.95	1.26	2.67	1.01	0.38	14
Other Race/ethnicity	0.88	1.06	1.27	1.96	1.08	0.35	19
Other Metro	1.11	1.33	1.56	2.47	1.36	0.37	26
Disadvantaged Urban	0.86	1.08	1.32	2.04	1.10	0.32	24
Advantaged Urban	0.86	1.17	1.49	2.61	1.20	0.41	17
Par. Ed. < HS	0.85	1.07	1.30	2.06	1.10	0.34	21
Par. Ed. = HS	0.97	1.17	1.35	2.20	1.21	0.31	29
Par. Ed. > HS	1.05	1.23	1.42	2.13	1.25	0.33	28
Par. Ed. = Coll.	1.16	1.37	1.64	2.46	1.41	0.36	30
Par. Ed. = IDK	0.89	1.06	1.24	1.74	1.05	0.28	29
Public School	1.04	1.24	1.39	2.37	1.25	0.30	34
Nonpublic School	1.06	1.36	1.70	2.79	1.43	0.50	16

---

* Distributions are based on 95 items.

Table D-7

Distributions of Design Effects by Demographic Subgroups for Proportion-correct Statistics Across Cognitive Mathematics Items in the Main Samples

Grade 4*

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.21	1.42	1.62	3.14	1.46	0.36	33
Male	1.13	1.29	1.51	2.12	1.33	0.28	45
Female	1.19	1.41	1.64	2.58	1.44	0.34	37
White	1.16	1.37	1.58	2.78	1.41	0.36	31
Black	1.02	1.23	1.48	2.83	1.27	0.33	29
Hispanic	1.08	1.32	1.49	2.36	1.33	0.33	33
Asian American	0.81	1.03	1.21	2.12	1.04	0.30	23
Other Race/ethnicity	0.93	1.14	1.36	2.85	1.16	0.36	20
Other Metro	1.20	1.43	1.66	3.57	1.46	0.39	28
Disadvantaged Urban	1.06	1.35	1.64	3.92	1.41	0.50	16
Advantaged Urban	1.01	1.34	1.69	3.86	1.42	0.50	16
Par. Ed. < HS	1.10	1.30	1.54	3.08	1.35	0.39	24
Par. Ed. = HS	1.09	1.28	1.56	3.08	1.35	0.37	26
Par. Ed. > HS	1.06	1.24	1.45	2.28	1.26	0.27	44
Par. Ed. = Coll.	1.18	1.37	1.55	2.29	1.38	0.31	40
Par. Ed. = IDK	1.15	1.37	1.64	2.44	1.40	0.34	34
Public School	1.18	1.36	1.60	2.91	1.42	0.34	35
Nonpublic School	0.99	1.14	1.43	2.57	1.22	0.37	22

---

* Distributions are based on 158 items.

Table D-8

Distributions of Design Effects by Demographic Subgroups for Proportion-correct Statistics Across Cognitive Mathematics Items in the Main Samples

Grade 8*

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.24	1.45	1.71	2.68	1.50	0.38	31
Male	1.18	1.43	1.63	2.56	1.44	0.33	39
Female	1.18	1.39	1.62	2.27	1.40	0.33	37
White	1.19	1.35	1.68	2.97	1.44	0.41	24
Black	1.16	1.41	1.68	2.80	1.44	0.40	26
Hispanic	1.05	1.21	1.47	2.32	1.28	0.35	27
Asian American	1.15	1.39	1.74	3.93	1.49	0.55	15
Other Race/ethnicity	0.79	1.05	1.34	3.20	1.10	0.45	12
Other Metro	1.14	1.39	1.68	3.14	1.47	0.43	23
Disadvantaged Urban	1.12	1.40	1.72	2.93	1.45	0.47	19
Advantaged Urban	1.18	1.53	1.99	3.92	1.64	0.65	13
Par. Ed. < HS	1.04	1.24	1.60	3.81	1.36	0.50	15
Par. Ed. = HS	1.09	1.31	1.52	2.43	1.32	0.31	35
Par. Ed. > HS	1.05	1.28	1.54	2.93	1.32	0.38	24
Par. Ed. = Coll.	1.21	1.44	1.69	2.98	1.48	0.35	36
Par. Ed. = IDK	1.09	1.33	1.56	3.15	1.37	0.39	25
Public School	1.17	1.37	1.63	2.75	1.42	0.37	29
Nonpublic School	1.06	1.24	1.49	2.90	1.32	0.37	25

---

* Distributions are based on 187 items.

Table D-9

Distributions of Design Effects by Demographic Subgroups for Proportion-correct Statistics Across Cognitive Mathematics Items in the Main Samples

Grade 12*

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.25	1.51	1.75	2.78	1.54	0.38	32
Male	1.19	1.39	1.62	2.47	1.43	0.34	36
Female	1.23	1.46	1.73	2.94	1.49	0.35	36
White	1.20	1.41	1.74	2.54	1.46	0.37	32
Black	1.17	1.38	1.64	4.10	1.43	0.42	23
Hispanic	1.16	1.47	1.98	4.25	1.66	0.75	10
Asian American	1.23	1.50	1.73	4.19	1.55	0.50	19
Other Race/ethnicity	0.92	1.30	1.74	4.28	1.47	0.81	7
Other Metro	1.19	1.45	1.74	2.68	1.51	0.39	29
Disadvantaged Urban	1.01	1.32	1.58	3.85	1.34	0.47	16
Advantaged Urban	1.13	1.42	1.95	4.13	1.57	0.59	14
Par. Ed. < HS	1.11	1.28	1.58	2.52	1.34	0.34	31
Par. Ed. = HS	1.10	1.31	1.54	2.08	1.31	0.30	37
Par. Ed. > HS	1.13	1.38	1.62	2.40	1.41	0.36	31
Par. Ed. = Coll.	1.25	1.48	1.72	2.28	1.48	0.33	39
Par. Ed. = IDK	1.17	1.46	1.62	3.33	1.43	0.45	20
Public School	1.21	1.42	1.65	2.86	1.45	0.37	30
Nonpublic School	1.23	1.54	1.89	5.24	1.65	0.68	12

---

* Distributions are based on 181 items.

Table D-10

Distributions of Design Effects by Demographic Subgroups for Proportion-correct Statistics Across Cognitive Mathematics Items in the Calculator Bridge Sample

Age 9*

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.36	1.78	1.99	2.85	1.74	0.45	29
Male	1.19	1.42	1.71	2.33	1.46	0.38	29
Female	1.27	1.48	1.72	2.39	1.50	0.36	34
White	1.25	1.64	2.00	2.56	1.65	0.45	26
Black	0.97	1.19	1.62	2.16	1.29	0.45	16
Hispanic	1.06	1.35	1.62	2.41	1.34	0.40	21
Asian American	0.62	0.81	1.05	2.37	0.93	0.44	9
Other Race/ethnicity	1.08	1.26	1.41	1.76	1.24	0.26	45
Other Metro	1.45	1.75	2.18	2.72	1.79	0.47	29
Disadvantaged Urban	0.73	1.08	1.61	4.59	1.30	0.84	5
Advantaged Urban	0.72	1.13	1.52	3.41	1.21	0.62	7
Par. Ed. < HS	1.01	1.15	1.24	1.94	1.17	0.27	37
Par. Ed. = HS	1.02	1.22	1.31	1.79	1.20	0.22	56
Par. Ed. > HS	1.16	1.36	1.52	1.96	1.34	0.30	39
Par. Ed. = Coll.	1.14	1.35	1.69	3.24	1.44	0.53	15
Par. Ed. = IDK	1.18	1.40	1.68	2.24	1.44	0.36	30
Public School	1.41	1.76	2.06	3.13	1.73	0.49	24
Nonpublic School	1.37	2.22	2.84	5.89	2.33	1.18	8

---

* Distributions are based on 38 items.

Table D-11

Distributions of Design Effects by Demographic Subgroups for Proportion-correct Statistics Across Cognitive Mathematics Items in the Main Estimation Samples

Grade 4*

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.35	1.66	2.08	3.19	1.74	0.45	29
Male	1.19	1.43	1.76	2.85	1.51	0.41	27
Female	1.16	1.42	1.86	2.80	1.54	0.44	23
White	1.36	1.59	2.06	3.29	1.70	0.52	21
Black	1.10	1.34	1.47	2.26	1.29	0.34	28
Hispanic	0.85	1.28	1.66	2.66	1.30	0.53	12
Asian American	1.09	1.41	1.59	3.37	1.38	0.47	17
Other Race/ethnicity	0.86	1.17	1.49	2.25	1.19	0.40	17
Other Metro	1.40	1.52	2.14	3.02	1.73	0.48	25
Disadvantaged Urban	1.15	1.65	2.30	4.68	1.82	0.89	8
Advantaged Urban	1.15	1.68	2.08	4.34	1.86	1.00	7
Par. Ed. < HS	0.90	1.08	1.21	1.89	1.08	0.25	35
Par. Ed. = HS	0.90	1.38	1.49	2.64	1.31	0.45	17
Par. Ed. > HS	0.87	1.19	1.36	1.77	1.17	0.34	23
Par. Ed. = Coll.	1.11	1.40	1.72	2.26	1.43	0.43	22
Par. Ed. = IDK	1.28	1.51	1.69	2.32	1.55	0.33	43
Public School	1.34	1.61	2.00	3.33	1.72	0.47	26
Nonpublic School	1.12	1.27	1.73	4.50	1.52	0.77	8

---

* Distributions are based on 33 items.

Table D-12

Distributions of Design Effects by Demographic Subgroups for Proportion-correct Statistics Across Cognitive Mathematics Items in the Main Estimation Samples

Grade 8*

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.23	1.64	1.98	3.81	1.74	0.64	14
Male	1.17	1.43	1.77	2.66	1.48	0.44	22
Female	1.13	1.38	1.59	2.83	1.43	0.44	21
White	1.05	1.57	2.09	3.71	1.62	0.69	11
Black	1.21	1.58	1.95	3.44	1.59	0.55	17
Hispanic	0.93	1.03	1.29	2.17	1.12	0.29	29
Asian American	0.74	0.93	1.15	2.36	1.01	0.34	17
Other Race/ethnicity	0.84	0.97	1.46	1.71	1.06	0.36	17
Other Metro	1.36	1.60	2.10	4.60	1.74	0.68	13
Disadvantaged Urban	1.12	2.10	2.88	5.82	2.12	1.08	8
Advantaged Urban	0.68	1.15	1.60	4.53	1.25	0.75	5
Par. Ed. < HS	0.90	1.05	1.18	1.74	1.05	0.27	31
Par. Ed. = HS	0.95	1.15	1.50	2.86	1.26	0.44	16
Par. Ed. > HS	1.12	1.36	1.63	2.43	1.40	0.35	32
Par. Ed. = Coll.	0.99	1.16	1.69	3.12	1.44	0.63	10
Par. Ed. = IDK	1.00	1.12	1.40	1.59	1.17	0.23	52
Public School	1.24	1.67	2.00	3.73	1.73	0.64	14
Nonpublic School	0.97	1.73	2.45	4.43	1.87	1.04	6

---

* Distributions are based on 42 items.

Table D-13

Distributions of Design Effects by Demographic Subgroups for Proportion-correct Statistics Across Cognitive Mathematics Items in the Main Estimation Samples

Grade 12*

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.16	1.46	1.70	2.54	1.50	0.39	29
Male	1.12	1.35	1.58	1.92	1.36	0.30	39
Female	1.07	1.21	1.43	1.99	1.28	0.26	47
White	1.22	1.40	1.60	2.72	1.46	0.42	23
Black	0.84	1.08	1.43	2.22	1.18	0.41	16
Hispanic	1.08	1.35	1.54	3.58	1.44	0.53	15
Asian American	1.13	1.54	1.89	4.28	1.64	0.70	11
Other Race/ethnicity	0.93	1.51	2.09	3.09	1.55	0.63	12
Other Metro	1.10	1.42	1.67	2.89	1.47	0.45	20
Disadvantaged Urban	0.93	1.19	1.59	2.81	1.31	0.55	11
Advantaged Urban	0.94	1.28	1.58	3.35	1.34	0.63	9
Par. Ed. < HS	0.92	1.08	1.45	2.54	1.23	0.45	15
Par. Ed. = HS	0.96	1.11	1.34	1.93	1.15	0.29	30
Par. Ed. > HS	0.95	1.11	1.31	1.73	1.13	0.26	36
Par. Ed. = Coll.	1.13	1.33	1.50	2.08	1.34	0.28	46
Par. Ed. = IDK	0.86	1.02	1.43	2.44	1.14	0.41	15
Public School	1.05	1.32	1.58	2.32	1.37	0.37	26
Nonpublic School	1.20	1.59	2.13	3.05	1.66	0.61	15

---

* Distributions are based on 44 items.

Table D-14

Distributions of Design Effects by Demographic Subgroups for Proportion-correct Statistics Across Cognitive Mathematics Items in the Long-term Trend Samples

Age 9*

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.44	1.77	2.26	4.40	1.93	0.72	14
Male	1.27	1.40	1.72	2.92	1.53	0.40	28
Female	1.25	1.53	1.76	4.08	1.64	0.61	14
White	1.26	1.43	1.83	3.24	1.61	0.53	18
Black	0.98	1.24	1.54	3.41	1.34	0.52	13
Hispanic	1.07	1.34	1.64	2.84	1.41	0.46	19
Asian American	0.90	1.20	1.63	4.02	1.40	0.65	9
Other Race/ethnicity	0.97	1.11	1.31	2.20	1.16	0.31	28
Other Metro	1.40	1.81	2.42	4.18	1.94	0.77	13
Disadvantaged Urban	1.09	1.53	2.01	4.13	1.60	0.73	9
Advantaged Urban	1.04	1.64	2.37	4.50	1.76	0.97	7
Par. Ed. < HS	0.95	1.15	1.39	2.33	1.21	0.32	28
Par. Ed. = HS	0.99	1.16	1.42	3.73	1.25	0.46	15
Par. Ed. > HS	0.98	1.16	1.31	1.88	1.14	0.28	33
Par. Ed. = Coll.	1.07	1.33	1.53	2.23	1.32	0.34	30
Par. Ed. = IDK	1.29	1.49	1.70	3.30	1.55	0.44	25
Public School	1.43	1.90	2.45	4.26	2.00	0.83	12
Nonpublic School	1.13	1.65	2.28	4.57	1.83	0.86	9

---

* Distributions are based on 68 items.

Table D-15

Distributions of Design Effects by Demographic Subgroups for Proportion-correct Statistics Across Cognitive Mathematics Items in the Long-term Trend Samples

Age 13*

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.37	1.68	2.22	4.42	1.83	0.68	14
Male	1.15	1.42	1.79	4.01	1.50	0.52	16
Female	1.16	1.53	1.99	2.89	1.57	0.50	19
White	1.27	1.58	1.92	4.35	1.69	0.62	14
Black	1.06	1.36	1.74	3.87	1.41	0.52	15
Hispanic	0.85	0.99	1.35	2.31	1.10	0.38	17
Asian American	0.79	1.04	1.33	2.88	1.11	0.46	11
Other Race/ethnicity	0.80	0.96	1.27	2.14	1.07	0.37	16
Other Metro	1.23	1.59	2.13	3.65	1.72	0.63	15
Disadvantaged Urban	1.08	1.62	2.38	6.69	1.86	1.13	5
Advantaged Urban	1.24	1.58	2.33	4.41	1.81	0.84	9
Par. Ed. < HS	0.85	1.02	1.23	2.13	1.07	0.32	22
Par. Ed. = HS	0.99	1.28	1.52	2.86	1.29	0.39	22
Par. Ed. > HS	0.97	1.16	1.46	2.18	1.22	0.36	22
Par. Ed. = Coll.	1.15	1.49	1.83	2.59	1.55	0.46	22
Par. Ed. = IDK	0.96	1.17	1.46	2.24	1.21	0.33	26
Public School	1.35	1.66	2.21	3.54	1.80	0.63	16
Nonpublic School	1.47	1.91	2.52	5.34	2.13	1.03	8

---

* Distributions are based on 98 items.

Table D-16

Distributions of Design Effects by Demographic Subgroups
for Proportion-correct Statistics
Across Cognitive Mathematics Items in the Long-term Trend Samples

Age 17*

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.39	1.89	2.30	4.13	1.94	0.66	17
Male	1.28	1.64	2.08	3.56	1.76	0.59	17
Female	1.21	1.56	1.88	2.61	1.56	0.43	26
White	1.31	1.57	1.95	3.71	1.65	0.58	16
Black	1.15	1.57	2.07	4.24	1.67	0.68	12
Hispanic	0.97	1.18	1.50	3.97	1.32	0.56	11
Asian American	0.82	1.05	1.25	2.70	1.09	0.39	16
Other Race/ethnicity	0.84	0.95	1.27	2.84	1.10	0.44	12
Other Metro	1.37	1.78	2.44	4.63	1.93	0.78	12
Disadvantaged Urban	1.01	1.36	1.97	3.36	1.52	0.73	9
Advantaged Urban	1.16	1.69	2.15	5.16	1.80	0.83	9
Par. Ed. < HS	0.94	1.30	1.78	2.89	1.40	0.56	12
Par. Ed. = HS	1.19	1.46	1.71	2.35	1.45	0.38	29
Par. Ed. > HS	1.19	1.48	1.81	4.05	1.53	0.47	21
Par. Ed. = Coll.	1.20	1.44	1.68	3.14	1.51	0.42	25
Par. Ed. = IDK	0.86	1.03	1.21	1.87	1.05	0.28	28
Public School	1.41	1.74	2.26	3.95	1.87	0.64	17
Nonpublic School	1.37	1.76	2.59	8.31	2.19	1.34	5

---

* Distributions are based on 94 items.

Table D-17

Distributions of Design Effects by Demographic Subgroups
for Proportion-correct Statistics
Across Cognitive Writing Items in the Main Samples

Grade 4*

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.80	1.97	2.17	2.27	1.86	0.43	****
Male	1.23	1.85	2.13	2.24	1.72	0.44	****
Female	1.43	1.48	1.57	2.52	1.55	0.41	****
White	1.72	1.92	1.99	2.01	1.74	0.35	****
Black	1.24	1.54	1.65	1.95	1.52	0.25	****
Hispanic	1.03	1.09	1.33	1.57	1.18	0.22	****
Asian American	0.86	0.95	1.02	1.94	1.02	0.37	****
Other Race/ethnicity	1.24	1.37	1.56	1.73	1.40	0.21	****
Other Metro	1.85	2.07	2.27	2.50	1.95	0.48	****
Disadvantaged Urban	1.21	1.44	1.80	1.92	1.45	0.38	****
Advantaged Urban	1.39	1.42	1.87	2.69	1.63	0.51	****
Par. Ed. < HS	1.01	1.22	1.82	2.14	1.38	0.48	****
Par. Ed. = HS	0.97	1.31	1.69	2.01	1.35	0.45	****
Par. Ed. > HS	1.32	1.42	1.58	1.82	1.45	0.23	****
Par. Ed. = Coll.	1.17	1.33	1.62	1.86	1.41	0.25	****
Par. Ed. = IDK	1.41	1.67	1.81	1.92	1.61	0.26	****
Public School	1.78	1.88	2.00	2.25	1.78	0.36	****
Nonpublic School	1.20	1.33	1.70	2.23	1.45	0.40	****

---

* Distributions are based on 9 items.

Table D-18

Distributions of Design Effects by Demographic Subgroups for Proportion-correct Statistics Across Cognitive Writing Items in the Main Samples

Grade 8*

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.01	1.16	1.60	1.85	1.28	0.32	****
Male	0.91	1.15	1.26	1.55	1.10	0.24	****
Female	1.18	1.25	1.52	1.69	1.33	0.19	****
White	0.98	1.14	1.46	1.97	1.24	0.35	****
Black	1.10	1.15	1.32	1.71	1.21	0.26	****
Hispanic	0.92	1.09	1.56	2.09	1.22	0.36	****
Asian American	0.97	1.14	1.42	1.88	1.21	0.33	****
Other Race/ethnicity	1.01	1.13	1.27	1.78	1.17	0.24	****
Other Metro	0.96	1.22	1.57	2.35	1.30	0.40	****
Disadvantaged Urban	0.92	1.37	1.65	1.80	1.29	0.37	****
Advantaged Urban	1.20	1.41	2.07	2.18	1.61	0.45	****
Par. Ed. < HS	0.95	0.99	1.24	1.45	1.09	0.20	****
Par. Ed. = HS	0.96	1.27	1.57	1.97	1.28	0.32	****
Par. Ed. > HS	0.90	1.06	1.39	1.57	1.14	0.28	****
Par. Ed. = Coll.	1.02	1.14	1.43	1.91	1.26	0.31	****
Par. Ed. = IDK	0.96	1.32	1.54	1.91	1.27	0.35	****
Public School	1.02	1.21	1.43	1.71	1.24	0.26	****
Nonpublic School	0.96	1.18	1.74	6.61	1.77	1.58	****

---

* Distributions are based on 11 items.

Table D-19

Distributions of Design Effects by Demographic Subgroups for Proportion-correct Statistics Across Cognitive Writing Items in the Main Samples

Grade 12*

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.47	1.71	1.82	2.35	1.68	0.29	****
Male	1.30	1.48	1.65	2.22	1.53	0.28	****
Female	1.30	1.43	1.73	2.71	1.54	0.42	****
White	1.51	1.75	2.10	2.31	1.76	0.34	****
Black	1.34	1.48	1.74	1.89	1.52	0.22	****
Hispanic	0.95	1.18	1.60	2.51	1.33	0.54	****
Asian American	1.15	1.36	1.48	2.01	1.37	0.29	****
Other Race/ethnicity	0.92	1.21	1.38	1.86	1.18	0.32	****
Other Metro	1.43	1.67	2.01	2.71	1.76	0.41	****
Disadvantaged Urban	1.09	1.24	1.54	1.93	1.29	0.34	****
Advantaged Urban	1.30	1.48	1.78	3.56	1.66	0.71	****
Par. Ed. < HS	1.06	1.15	1.29	1.88	1.21	0.25	****
Par. Ed. = HS	1.20	1.33	1.66	2.36	1.43	0.39	****
Par. Ed. > HS	1.10	1.21	1.53	2.21	1.35	0.43	****
Par. Ed. = Coll.	1.19	1.42	1.68	2.23	1.44	0.39	****
Par. Ed. = IDK	0.98	1.49	1.76	3.21	1.54	0.78	****
Public School	1.37	1.42	1.66	2.19	1.53	0.25	****
Nonpublic School	1.21	1.63	2.75	4.08	2.03	1.06	****

---

* Distributions are based on 12 items.

Table D-20

Distributions of Design Effects by Demographic Subgroups for Proportion-correct Statistics Across Cognitive Writing Items in the Long-term Trend Samples

Grade 4*

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.28	1.48	1.48	1.54	1.39	0.15	****
Male	1.08	1.16	1.29	1.41	1.20	0.13	****
Female	1.19	1.30	1.84	1.88	1.48	0.32	****
White	1.17	1.17	1.48	1.55	1.28	0.20	****
Black	0.95	1.41	1.61	1.71	1.29	0.36	****
Hispanic	1.23	1.24	1.50	1.74	1.27	0.37	****
Asian American	0.52	0.83	0.93	1.00	0.75	0.22	****
Other Race/ethnicity	0.74	0.78	0.87	1.32	0.86	0.24	****
Other Metro	1.15	1.47	1.52	1.64	1.37	0.23	****
Disadvantaged Urban	1.08	1.15	1.61	2.27	1.42	0.48	****
Advantaged Urban	1.17	1.19	1.39	1.44	1.16	0.30	****
Par. Ed. < HS	0.93	1.22	1.26	1.88	1.24	0.35	****
Par. Ed. = HS	0.98	1.61	1.74	1.83	1.39	0.42	****
Par. Ed. > HS	1.21	1.29	1.34	1.73	1.33	0.22	****
Par. Ed. = Coll.	1.12	1.15	1.57	1.67	1.31	0.26	****
Par. Ed. = IDK	1.17	1.19	1.28	1.34	1.16	0.19	****
Public School	1.27	1.50	1.70	1.70	1.43	0.27	****
Nonpublic School	0.96	1.02	1.14	1.69	1.07	0.37	****

---

* Distributions are based on 5 items.

Table D-21

Distributions of Design Effects by Demographic Subgroups for Proportion-correct Statistics Across Cognitive Writing Items in the Long-term Trend Samples

Grade 8*

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.35	1.46	2.01	2.29	1.61	0.41	****
Male	0.85	1.19	1.70	1.94	1.26	0.46	****
Female	1.11	1.18	1.39	1.61	1.25	0.20	****
White	1.19	1.45	1.60	1.69	1.42	0.21	****
Black	1.34	1.51	1.63	1.92	1.45	0.34	****
Hispanic	0.99	1.15	1.52	1.87	1.27	0.34	****
Asian American	0.99	1.40	1.86	2.18	1.44	0.52	****
Other Race/ethnicity	0.70	1.11	1.29	1.41	1.03	0.31	****
Other Metro	1.41	1.65	1.96	2.20	1.62	0.44	****
Disadvantaged Urban	0.80	1.25	1.50	1.60	1.20	0.33	****
Advantaged Urban	0.94	2.25	3.84	4.56	2.42	1.43	****
Par. Ed. < HS	0.64	0.97	1.15	1.22	0.91	0.26	****
Par. Ed. = HS	1.13	1.16	1.38	1.69	1.26	0.22	****
Par. Ed. > HS	1.02	1.36	1.56	1.86	1.35	0.32	****
Par. Ed. = Coll.	0.94	1.48	1.69	1.73	1.34	0.39	****
Par. Ed. = IDK	1.16	1.28	1.46	1.48	1.25	0.21	****
Public School	1.11	1.34	1.77	1.78	1.36	0.34	****
Nonpublic School	1.45	1.86	2.02	3.49	1.97	0.74	****

---

* Distributions are based on 6 items.

Table D-22

Distributions of Design Effects by Demographic Subgroups for Proportion-correct Statistics Across Cognitive Writing Items in the Long-term Trend Samples

Grade 12*

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.31	1.64	1.93	2.01	1.62	0.30	****
Male	1.35	1.52	1.63	1.69	1.43	0.27	****
Female	1.21	1.34	1.51	1.62	1.36	0.17	****
White	1.39	1.62	1.70	2.06	1.63	0.23	****
Black	1.16	1.28	1.58	1.83	1.36	0.27	****
Hispanic	0.72	0.97	1.35	1.41	1.01	0.31	****
Asian American	0.99	1.28	2.00	2.99	1.57	0.74	****
Other Race/ethnicity	1.02	1.13	1.40	1.81	1.21	0.33	****
Other Metro	1.48	1.67	1.94	2.06	1.70	0.24	****
Disadvantaged Urban	1.01	1.39	2.01	2.20	1.47	0.50	****
Advantaged Urban	0.89	1.22	1.55	1.82	1.25	0.35	****
Par. Ed. < HS	1.01	1.14	1.20	1.33	1.10	0.18	****
Par. Ed. = HS	1.21	1.30	1.50	1.63	1.32	0.21	****
Par. Ed. > HS	1.28	1.48	1.79	2.14	1.56	0.33	****
Par. Ed. = Coll.	1.01	1.36	1.60	1.70	1.34	0.27	****
Par. Ed. = IDK	0.71	0.92	1.50	1.89	1.09	0.46	****
Public School	1.21	1.70	1.89	1.98	1.57	0.37	****
Nonpublic School	1.27	1.80	2.25	2.41	1.78	0.48	****

---

* Distributions are based on 6 items.

Table D-23

Distributions of Design Effects by Demographic Subgroups for Proportion-correct Statistics Across Cognitive Science Items in the Long-term Trend Samples

Age 9*

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.29	1.67	2.17	3.75	1.74	0.58	18
Male	1.08	1.33	1.64	2.94	1.42	0.46	18
Female	1.15	1.44	1.70	2.85	1.48	0.42	24
White	1.16	1.48	2.00	3.28	1.61	0.55	17
Black	0.87	1.15	1.54	2.20	1.22	0.41	17
Hispanic	0.96	1.21	1.49	2.61	1.26	0.40	19
Asian American	0.83	1.09	1.40	2.40	1.15	0.39	17
Other Race/ethnicity	0.93	1.09	1.45	3.17	1.27	0.53	11
Other Metro	1.32	1.65	2.23	3.97	1.77	0.63	16
Disadvantaged Urban	0.84	1.19	1.77	3.93	1.36	0.77	6
Advantaged Urban	0.93	1.29	1.60	2.58	1.30	0.47	15
Par. Ed. < HS	0.87	1.13	1.34	2.06	1.15	0.35	22
Par. Ed. = HS	0.95	1.16	1.44	2.57	1.20	0.35	23
Par. Ed. > HS	0.88	1.05	1.23	1.92	1.10	0.31	24
Par. Ed. = Coll.	1.02	1.23	1.40	2.81	1.26	0.37	23
Par. Ed. = IDK	1.03	1.34	1.73	2.52	1.42	0.44	21
Public School	1.29	1.74	2.08	3.46	1.72	0.55	19
Nonpublic School	1.06	1.34	1.84	3.93	1.53	0.68	10

---

* Distributions are based on 63 items.

Table D-24

Distributions of Design Effects by Demographic Subgroups for Proportion-correct Statistics Across Cognitive Science Items in the Long-term Trend Samples

Age 13*

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.28	1.57	1.90	2.61	1.60	0.42	29
Male	1.10	1.27	1.59	2.39	1.36	0.37	27
Female	1.15	1.34	1.56	2.14	1.38	0.35	31
White	1.20	1.47	1.76	2.77	1.50	0.40	28
Black	1.00	1.24	1.50	2.74	1.30	0.42	19
Hispanic	0.86	1.03	1.26	2.01	1.08	0.35	19
Asian American	0.73	1.00	1.29	2.93	1.07	0.47	10
Other Race/ethnicity	0.67	0.97	1.25	3.22	1.03	0.46	10
Other Metro	1.23	1.47	1.80	2.66	1.51	0.40	29
Disadvantaged Urban	0.92	1.47	2.08	5.79	1.65	0.93	6
Advantaged Urban	1.07	1.43	2.20	4.62	1.70	0.83	8
Par. Ed. < HS	0.83	1.06	1.37	2.00	1.10	0.34	21
Par. Ed. = HS	0.98	1.19	1.42	2.24	1.24	0.35	26
Par. Ed. > HS	0.91	1.11	1.32	1.98	1.15	0.30	29
Par. Ed. = Coll.	1.09	1.29	1.55	2.06	1.32	0.32	34
Par. Ed. = IDK	0.91	1.09	1.30	1.81	1.10	0.27	32
Public School	1.22	1.53	1.84	2.91	1.57	0.44	25
Nonpublic School	1.18	1.61	2.12	3.45	1.71	0.62	15

---

* Distributions are based on 83 items.

Table D-25

Distributions of Design Effects by Demographic Subgroups for Proportion-correct Statistics Across Cognitive Science Items in the Long-term Trend Samples

Age 17[*]

Subgroup	Bottom Quartile	Median	Top Quartile	Max.	Mean	Standard Deviation	Degrees of Freedom
Total	1.40	1.66	1.99	2.91	1.70	0.45	28
Male	1.29	1.52	1.76	2.77	1.53	0.35	37
Female	1.25	1.47	1.67	2.68	1.50	0.42	25
White	1.25	1.55	1.80	2.63	1.55	0.42	27
Black	1.05	1.34	1.74	4.14	1.44	0.62	11
Hispanic	1.04	1.23	1.73	3.69	1.47	0.70	9
Asian American	0.94	1.16	1.45	2.00	1.18	0.35	23
Other Race/ethnicity	0.89	1.17	1.43	2.79	1.18	0.44	14
Other Metro	1.38	1.70	2.17	3.51	1.79	0.58	19
Disadvantaged Urban	0.89	1.17	1.46	3.13	1.26	0.53	11
Advantaged Urban	1.20	1.57	2.01	3.09	1.64	0.63	13
Par. Ed. < HS	1.13	1.45	1.76	3.27	1.47	0.50	17
Par. Ed. = HS	1.12	1.38	1.63	2.14	1.39	0.34	34
Par. Ed. > HS	1.13	1.31	1.78	2.64	1.41	0.42	23
Par. Ed. = Coll.	1.05	1.26	1.56	3.04	1.32	0.39	23
Par. Ed. = IDK	0.86	1.12	1.34	1.85	1.14	0.31	26
Public School	1.34	1.62	1.93	3.03	1.68	0.44	28
Nonpublic School	1.14	1.55	2.16	4.04	1.67	0.76	9

---

[*] Distributions are based on 82 items.

APPENDIX E

Subscale Item Counts and IRT Parameters

APPENDIX E

IRT Parameters

This appendix contains tables of IRT (item response theory) parameters for NAEP items that were scaled in each subject area and study (main and long-term trend) for which IRT scales were created.

For each of the binary scored items used in scaling (i.e., multiple-choice items and short constructed-response items), the tables provide estimates of the IRT parameters (which correspond to a_j, b_j, and c_j in equation 11.1 in Chapter 11) and their associated standard errors (s.e.) of the estimates. For each of the polytomously scored items (i.e., the extended constructed-response items and the testlets), the tables also show the estimates of the d_{jv} parameters (see equation 11.1) and their associated standard errors.

The tables also show the block in which each item appears for each age class (*Block*) and the position of each item within its block (*Item*).

Note that item parameters shown in this appendix are in the metrics used for the original calibration of the scales. The transformations needed to represent these parameters in terms of the metric of the final reporting scales are given in Chapters 12 through 15.

Table E-1
IRT Parameters for Reading Long-term Trend Samples
Age 9

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
N001101	0.564	(0.165)	2.517	(0.416)	0.427	(0.023)			84H	5
N001501	1.960	(0.156)	-0.623	(0.053)	0.360	(0.026)			84H	10
N001502	2.138	(0.149)	0.310	(0.029)	0.203	(0.015)			84H	11
N001503	1.655	(0.127)	-0.156	(0.051)	0.333	(0.023)			84H	12
N001504	2.499	(0.207)	0.290	(0.032)	0.341	(0.017)			84H	13
N001507	0.589	(0.102)	3.259	(0.467)	0.000	(0.000)			84H	15
N001601	0.892	(0.078)	0.160	(0.077)	0.260	(0.027)			84J	12
N001602	1.606	(0.142)	0.573	(0.041)	0.295	(0.018)			84J	13
N001603	1.005	(0.130)	1.079	(0.069)	0.294	(0.020)			84J	14
N001604	1.428	(0.148)	0.954	(0.044)	0.254	(0.016)			84J	15
N001802	2.381	(0.202)	1.781	(0.068)	0.219	(0.009)			84J	20
N002001	1.418	(0.140)	1.020	(0.042)	0.190	(0.015)			84K	9
N002002	1.189	(0.114)	0.847	(0.048)	0.217	(0.017)			84K	10
N002003	1.257	(0.105)	0.539	(0.046)	0.224	(0.019)			84K	11
N002101	1.568	(0.197)	1.884	(0.108)	0.237	(0.011)			84K	18
N002102	1.433	(0.235)	2.250	(0.176)	0.154	(0.010)			84K	19
N002401	1.489	(0.114)	0.678	(0.034)	0.145	(0.014)			84L	22
N002702	1.462	(0.134)	0.909	(0.039)	0.185	(0.015)			84L	20
N002801	2.234	(0.162)	0.229	(0.031)	0.209	(0.017)			84L	24
N002802	1.764	(0.132)	0.058	(0.042)	0.242	(0.021)			84L	25
N002804	0.545	(0.050)	2.065	(0.155)	0.000	(0.000)			84L	26
N003001	0.849	(0.159)	2.024	(0.173)	0.165	(0.016)			84M	10
N003002	0.282	(0.046)	1.148	(0.288)	0.218	(0.046)			84M	11
N003101	1.123	(0.100)	0.243	(0.064)	0.276	(0.024)			84M	14
N003102	1.651	(0.148)	0.765	(0.037)	0.203	(0.015)			84M	15
N003104	0.571	(0.085)	3.474	(0.436)	0.000	(0.000)			84M	16
N003701	1.060	(0.103)	0.072	(0.083)	0.369	(0.028)			84N	23
N003702	1.412	(0.129)	0.578	(0.047)	0.266	(0.019)			84N	24
N003704	0.747	(0.068)	0.934	(0.077)	0.000	(0.000)			84N	25
N003801	0.721	(0.158)	1.998	(0.203)	0.281	(0.021)			84O	12
N003802	0.658	(0.073)	0.485	(0.102)	0.223	(0.032)			84O	13
N003803	0.764	(0.273)	3.455	(0.805)	0.228	(0.013)			84O	14
N004101	0.698	(0.064)	-0.513	(0.137)	0.285	(0.041)			84O	17
N004201	0.993	(0.131)	1.232	(0.072)	0.232	(0.019)			84O	18
N004202	0.740	(0.118)	1.218	(0.105)	0.311	(0.026)			84O	19
N004701	1.433	(0.108)	0.446	(0.040)	0.201	(0.017)			84Q	10
N004702	0.699	(0.078)	0.131	(0.125)	0.334	(0.035)			84Q	11
N004703	0.949	(0.079)	0.234	(0.066)	0.225	(0.024)			84Q	12
N004801	1.321	(0.106)	-0.260	(0.067)	0.360	(0.027)			84Q	13
N004901	1.589	(0.138)	1.018	(0.039)	0.222	(0.014)			84Q	14
N005101	0.537	(0.046)	-2.153	(0.250)	0.291	(0.063)			84Q	15
N008601	1.360	(0.092)	-0.395	(0.056)	0.237	(0.025)			84H	6
N008602	1.078	(0.080)	0.036	(0.060)	0.220	(0.024)			84H	7
N008603	1.014	(0.068)	-0.448	(0.068)	0.192	(0.026)			84H	8
N008701	0.545	(0.048)	-3.402	(0.282)	0.281	(0.064)			84H	9
N008801	1.225	(0.094)	-0.910	(0.085)	0.336	(0.033)			84J	18
N008901	1.788	(0.140)	-0.008	(0.045)	0.321	(0.021)			84J	21
N008902	1.151	(0.089)	-0.258	(0.070)	0.290	(0.027)			84J	22
N009001	1.820	(0.142)	0.596	(0.032)	0.201	(0.015)			84K	12
N009002	1.410	(0.139)	0.853	(0.044)	0.248	(0.017)			84K	13
N009003	1.585	(0.144)	1.354	(0.054)	0.253	(0.012)			84K	14
N009004	1.425	(0.113)	0.403	(0.043)	0.221	(0.019)			84K	15
N009101	1.005	(0.078)	-0.668	(0.092)	0.294	(0.033)			84K	16
N009201	1.510	(0.118)	-0.764	(0.068)	0.331	(0.030)			84K	17

Table E-1 (continued)
IRT Parameters for Reading Long-term Trend Samples
Age 9

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
N009401	1.222	(0.078)	-0.746	(0.063)	0.173	(0.027)			84L	23
N009601	0.703	(0.048)	-1.776	(0.142)	0.180	(0.045)			84L	21
N009701	1.310	(0.110)	0.295	(0.053)	0.295	(0.022)			84M	5
N009702	1.730	(0.132)	0.249	(0.040)	0.278	(0.019)			84M	6
N009703	1.333	(0.128)	0.746	(0.047)	0.265	(0.018)			84M	7
N009704	1.181	(0.106)	0.773	(0.046)	0.190	(0.017)			84M	8
N009705	1.662	(0.128)	0.181	(0.043)	0.287	(0.020)			84M	9
N009801	1.065	(0.089)	-1.841	(0.137)	0.320	(0.050)			84N	12
N009901	1.016	(0.086)	-0.018	(0.076)	0.306	(0.027)			84N	13
N010002	1.093	(0.078)	-0.249	(0.065)	0.225	(0.026)			84N	18
N010003	1.424	(0.109)	0.096	(0.050)	0.276	(0.022)			84N	19
N010102	1.101	(0.116)	0.696	(0.060)	0.292	(0.022)			84N	21
N010103	1.776	(0.128)	-0.048	(0.042)	0.250	(0.021)			84N	22
N010201	0.987	(0.080)	-1.353	(0.124)	0.314	(0.044)			84O	16
N010301	0.523	(0.051)	-1.301	(0.243)	0.307	(0.058)			84O	15
N010401	0.605	(0.062)	-0.721	(0.185)	0.304	(0.049)			84O	20
N010402	1.158	(0.140)	1.085	(0.059)	0.239	(0.018)			84O	21
N010403	1.419	(0.151)	1.386	(0.063)	0.235	(0.013)			84O	22
N010801	1.154	(0.095)	0.339	(0.054)	0.238	(0.022)			84Q	16
N010902	1.911	(0.164)	0.570	(0.036)	0.275	(0.017)			84Q	18
N010903	1.944	(0.147)	0.281	(0.035)	0.245	(0.018)			84Q	19
N010904	2.178	(0.178)	0.694	(0.032)	0.303	(0.015)			84Q	20
N011001	1.292	(0.072)	0.042	(0.041)	0.289	(0.017)			84R	5
N011002	1.889	(0.117)	0.640	(0.025)	0.276	(0.012)			84R	6
N011003	2.206	(0.113)	-0.093	(0.025)	0.267	(0.014)			84R	7
N011004	1.780	(0.095)	0.342	(0.026)	0.242	(0.013)			84R	8
N011101	1.592	(0.084)	0.335	(0.027)	0.210	(0.013)			84R	9
N011201	0.980	(0.071)	0.579	(0.047)	0.276	(0.017)			84R	10
N011301	1.900	(0.107)	0.255	(0.027)	0.299	(0.014)			84R	11
N011302	1.023	(0.077)	0.542	(0.050)	0.309	(0.018)			84R	12
N011401	1.776	(0.120)	1.295	(0.040)	0.370	(0.010)			84R	13
N011402	0.814	(0.088)	1.071	(0.066)	0.314	(0.019)			84R	14
N011403	1.657	(0.109)	1.295	(0.038)	0.277	(0.009)			84R	15
N011404	1.720	(0.105)	1.255	(0.034)	0.243	(0.009)			84R	16
N013201	1.467	(0.104)	0.107	(0.043)	0.223	(0.020)			84V	29
N013301	1.318	(0.104)	-0.647	(0.076)	0.377	(0.030)			84V	30
N013401	1.267	(0.104)	0.588	(0.044)	0.196	(0.018)			84V	31
N013402	1.524	(0.115)	0.100	(0.047)	0.289	(0.021)			84V	32
N013403	1.731	(0.143)	0.576	(0.036)	0.253	(0.016)			84V	33
N014001	1.201	(0.095)	-0.094	(0.064)	0.290	(0.026)			84M	13
N014101	0.746	(0.070)	-0.196	(0.113)	0.259	(0.035)			84Q	21
N014201	0.874	(0.067)	-0.421	(0.089)	0.242	(0.031)			84V	34
N014301	2.089	(0.150)	0.222	(0.032)	0.227	(0.017)			84N	14
N014302	1.268	(0.113)	0.521	(0.050)	0.264	(0.020)			84N	15
N014303	1.819	(0.125)	-0.090	(0.039)	0.234	(0.020)			84N	16
N014501	0.526	(0.029)	-1.142	(0.074)	0.000	(0.000)			84V	35
N014502	0.514	(0.029)	-1.395	(0.083)	0.000	(0.000)			84V	35
N014503	0.744	(0.037)	-1.767	(0.074)	0.000	(0.000)			84V	35

BEST COPY AVAILABLE

Table E-2
IRT Parameters for Reading Long-term Trend Samples
Age 13

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
N001101	0.172	(0.035)	1.462	(0.533)	0.305	(0.044)			84H	6
N001201	0.816	(0.175)	1.625	(0.157)	0.422	(0.024)			84H	7
N001202	0.973	(0.123)	0.953	(0.076)	0.258	(0.024)			84H	8
N001301	0.662	(0.090)	0.179	(0.170)	0.408	(0.043)			84H	9
N001302	0.523	(0.068)	-2.713	(0.397)	0.502	(0.070)			84H	10
N001303	1.133	(0.132)	0.718	(0.068)	0.298	(0.024)			84H	11
N001401	0.819	(0.088)	0.061	(0.113)	0.286	(0.039)			84H	12
N001501	1.445	(0.132)	-1.923	(0.105)	0.233	(0.050)			84H	13
N001502	1.431	(0.107)	-0.830	(0.063)	0.193	(0.033)			84H	14
N001503	1.213	(0.109)	-1.261	(0.111)	0.294	(0.048)			84H	15
N001504	1.101	(0.086)	-0.936	(0.089)	0.225	(0.039)			84H	16
N001507	0.540	(0.071)	2.310	(0.256)	0.000	(0.000)			84H	18
N001601	0.388	(0.045)	-1.231	(0.303)	0.270	(0.061)			84J	11
N001602	0.806	(0.069)	-1.340	(0.149)	0.260	(0.053)			84J	12
N001603	0.541	(0.065)	-0.233	(0.196)	0.276	(0.052)			84J	13
N001604	1.083	(0.094)	-0.233	(0.082)	0.281	(0.034)			84J	14
N001701	0.839	(0.078)	-0.796	(0.136)	0.294	(0.048)			84J	17
N001702	0.395	(0.115)	3.513	(0.653)	0.224	(0.029)			84J	18
N001703	0.671	(0.066)	-0.213	(0.129)	0.220	(0.042)			84J	19
N001802	1.083	(0.118)	0.712	(0.065)	0.247	(0.024)			84J	21
N001901	0.893	(0.089)	0.093	(0.093)	0.246	(0.034)			84J	22
N002001	1.108	(0.085)	-0.127	(0.064)	0.186	(0.028)			84K	9
N002002	1.003	(0.079)	-0.323	(0.078)	0.205	(0.033)			84K	10
N002003	1.139	(0.089)	-0.647	(0.079)	0.233	(0.035)			84K	11
N002101	0.595	(0.087)	1.138	(0.124)	0.198	(0.033)			84K	12
N002102	1.098	(0.103)	0.769	(0.053)	0.136	(0.019)			84K	13
N002201	1.011	(0.091)	-0.246	(0.090)	0.271	(0.036)			84K	14
N002202	0.789	(0.080)	-0.818	(0.161)	0.321	(0.053)			84K	15
N002203	0.541	(0.050)	-1.840	(0.216)	0.243	(0.057)			84K	16
N002401	0.909	(0.068)	-0.960	(0.097)	0.162	(0.038)			84L	22
N002501	0.690	(0.074)	0.195	(0.116)	0.224	(0.038)			84L	23
N002701	1.106	(0.127)	0.628	(0.071)	0.311	(0.025)			84L	24
N002801	1.702	(0.151)	-1.073	(0.073)	0.272	(0.038)			84L	25
N002802	1.008	(0.084)	-1.640	(0.129)	0.244	(0.051)			84L	26
N002902	0.466	(0.052)	-1.587	(0.297)	0.301	(0.067)			84M	6
N002903	1.706	(0.141)	-0.670	(0.059)	0.289	(0.032)			84M	7
N002904	0.978	(0.082)	-0.232	(0.083)	0.232	(0.034)			84M	8
N002905	0.634	(0.069)	0.359	(0.117)	0.206	(0.036)			84M	9
N002906	1.533	(0.125)	-0.718	(0.066)	0.270	(0.034)			84M	10
N003001	0.722	(0.110)	1.400	(0.110)	0.189	(0.026)			84M	11
N003002	0.313	(0.039)	-0.230	(0.259)	0.173	(0.049)			84M	12
N003003	2.092	(0.183)	1.990	(0.077)	0.068	(0.006)			84M	13
N003101	0.985	(0.084)	-1.117	(0.119)	0.261	(0.048)			84M	14
N003102	1.451	(0.115)	-0.476	(0.063)	0.260	(0.032)			84M	15
N003104	0.492	(0.057)	1.438	(0.154)	0.000	(0.000)			84M	16
N003201	0.832	(0.076)	-0.810	(0.135)	0.287	(0.048)			84N	12
N003202	0.861	(0.073)	-0.019	(0.082)	0.179	(0.031)			84N	13
N003203	0.856	(0.078)	-0.040	(0.092)	0.221	(0.034)			84N	14
N003204	0.634	(0.082)	0.568	(0.129)	0.247	(0.038)			84N	15
N003301	0.839	(0.070)	-0.616	(0.107)	0.215	(0.040)			84N	16
N003401	1.014	(0.084)	-0.289	(0.081)	0.213	(0.034)			84N	17
N003501	0.696	(0.069)	-0.492	(0.146)	0.263	(0.047)			84N	18
N003601	1.116	(0.099)	-1.092	(0.112)	0.294	(0.047)			84N	19
N003602	0.897	(0.073)	-0.266	(0.084)	0.185	(0.033)			84N	20

Table E-2 (continued)
IRT Parameters for Reading Long-term Trend Samples
Age 13

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
N003701	1.033	(0.095)	-0.567	(0.103)	0.305	(0.041)			84N	21
N003702	1.178	(0.116)	0.019	(0.079)	0.319	(0.032)			84N	22
N003704	0.713	(0.047)	0.286	(0.050)	0.000	(0.000)			84N	23
N003801	0.495	(0.095)	1.429	(0.186)	0.265	(0.041)			84O	12
N003802	0.237	(0.034)	-1.494	(0.426)	0.202	(0.058)			84O	13
N003803	0.393	(0.105)	2.896	(0.463)	0.250	(0.034)			84O	14
N003901	0.983	(0.096)	-2.594	(0.175)	0.253	(0.059)			84O	16
N004002	0.410	(0.046)	-2.662	(0.347)	0.267	(0.063)			84O	15
N004101	0.788	(0.067)	-1.759	(0.161)	0.248	(0.056)			84O	17
N004201	0.766	(0.081)	-0.060	(0.124)	0.286	(0.041)			84O	18
N004202	0.538	(0.066)	-0.041	(0.187)	0.265	(0.050)			84O	19
N004301	1.248	(0.121)	0.270	(0.065)	0.291	(0.027)			84O	20
N004303	0.834	(0.050)	0.389	(0.043)	0.000	(0.000)			84O	21
N004401	0.838	(0.082)	-2.780	(0.204)	0.252	(0.060)			84P	7
N004402	0.842	(0.074)	-0.170	(0.094)	0.217	(0.035)			84P	8
N004403	1.197	(0.103)	-1.724	(0.118)	0.252	(0.052)			84P	9
N004501	0.676	(0.077)	0.304	(0.121)	0.244	(0.038)			84P	10
N004502	0.578	(0.056)	-1.004	(0.193)	0.257	(0.054)			84P	11
N004601	0.905	(0.097)	0.395	(0.086)	0.264	(0.031)			84P	12
N004602	1.024	(0.087)	-0.122	(0.077)	0.221	(0.032)			84P	13
N004603	1.552	(0.124)	-0.512	(0.059)	0.261	(0.031)			84P	14
N004605	0.786	(0.047)	-0.958	(0.063)	0.000	(0.000)			84P	15
N004701	1.333	(0.103)	-0.923	(0.073)	0.214	(0.035)			84Q	7
N004702	0.793	(0.067)	-1.194	(0.138)	0.241	(0.049)			84Q	8
N004703	0.725	(0.066)	-1.181	(0.162)	0.267	(0.053)			84Q	9
N004801	1.090	(0.089)	-1.356	(0.109)	0.239	(0.046)			84Q	10
N004901	0.931	(0.084)	-0.169	(0.091)	0.253	(0.035)			84Q	11
N005002	0.415	(0.100)	2.277	(0.325)	0.278	(0.038)			84Q	14
N005003	0.892	(0.150)	1.930	(0.141)	0.138	(0.016)			84Q	15
N005101	0.371	(0.052)	-4.239	(0.572)	0.269	(0.064)			84Q	12
N005201	0.691	(0.162)	1.253	(0.196)	0.576	(0.030)			84Q	16
N005202	0.410	(0.058)	0.524	(0.214)	0.229	(0.047)			84Q	17
N005203	0.871	(0.226)	2.358	(0.268)	0.309	(0.018)			84Q	18
N005301	1.126	(0.109)	0.029	(0.078)	0.282	(0.032)			84Q	19
N005302	1.605	(0.156)	0.550	(0.046)	0.214	(0.020)			84Q	20
N005303	0.551	(0.085)	0.928	(0.149)	0.217	(0.038)			84Q	21
N005304	2.200	(0.193)	0.133	(0.038)	0.220	(0.021)			84Q	22
N005305	1.430	(0.152)	-0.379	(0.085)	0.410	(0.035)			84Q	23
N005403	0.967	(0.085)	-0.707	(0.109)	0.280	(0.043)			84R	7
N005404	0.861	(0.077)	-1.868	(0.169)	0.277	(0.060)			84R	8
N005405	1.778	(0.151)	0.098	(0.046)	0.283	(0.024)			84R	9
N005406	0.869	(0.077)	-0.482	(0.108)	0.252	(0.041)			84R	10
N005407	1.097	(0.095)	-0.663	(0.096)	0.286	(0.041)			84R	11
N005503	0.968	(0.139)	0.702	(0.099)	0.420	(0.029)			84R	14
N005504	1.139	(0.138)	1.075	(0.064)	0.215	(0.020)			84R	15
N005505	0.629	(0.056)	-1.485	(0.184)	0.245	(0.055)			84R	16
N005601	1.480	(0.130)	-0.587	(0.073)	0.334	(0.036)			84R	17
N005602	1.369	(0.119)	0.522	(0.046)	0.184	(0.020)			84R	18
N005603	0.945	(0.083)	-0.662	(0.109)	0.269	(0.043)			84R	19

Table E-3
IRT Parameters for Reading Long-term Trend Samples
Age 17

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
N001301	0.709	(0.094)	-0.350	(0.192)	0.505	(0.044)			84BH	10
N001302	0.541	(0.071)	-3.073	(0.441)	0.563	(0.069)			84BH	11
N001303	1.008	(0.093)	-0.271	(0.093)	0.324	(0.035)			84BH	12
N001401	0.910	(0.100)	-0.424	(0.134)	0.453	(0.040)			84BH	13
N001501	1.344	(0.137)	-2.440	(0.141)	0.286	(0.055)			84BH	14
N001502	1.302	(0.102)	-1.305	(0.083)	0.190	(0.034)			84BH	15
N001503	1.151	(0.106)	-1.667	(0.130)	0.279	(0.047)			84BH	16
N001504	1.140	(0.094)	-1.455	(0.106)	0.233	(0.039)			84BH	17
N001507	0.601	(0.066)	1.272	(0.125)	0.000	(0.000)			84BH	19
N001701	0.692	(0.072)	-1.313	(0.203)	0.342	(0.056)			84BJ	12
N001703	0.653	(0.071)	-0.897	(0.191)	0.336	(0.052)			84BJ	14
N001901	0.957	(0.093)	-1.044	(0.135)	0.362	(0.046)			84BJ	15
N001904	0.747	(0.045)	-1.117	(0.070)	0.000	(0.000)			84BJ	17
N002001	0.947	(0.085)	-0.639	(0.107)	0.301	(0.040)			84BK	9
N002002	1.208	(0.106)	-0.767	(0.091)	0.338	(0.038)			84BK	10
N002003	1.192	(0.105)	-1.169	(0.103)	0.306	(0.042)			84BK	11
N002101	0.620	(0.084)	0.405	(0.141)	0.278	(0.041)			84BK	12
N002102	1.521	(0.126)	0.106	(0.050)	0.233	(0.025)			84BK	13
N002201	1.783	(0.168)	-0.961	(0.071)	0.376	(0.036)			84BK	14
N002202	2.622	(0.312)	-0.832	(0.060)	0.496	(0.034)			84BK	15
N002203	0.385	(0.049)	-3.777	(0.487)	0.320	(0.066)			84BK	16
N002501	0.527	(0.061)	-0.620	(0.211)	0.293	(0.052)			84BL	27
N002701	0.817	(0.077)	-0.155	(0.103)	0.253	(0.037)			84BL	28
N002702	0.750	(0.058)	-1.281	(0.125)	0.166	(0.040)			84BL	29
N002801	1.178	(0.115)	-2.336	(0.160)	0.300	(0.057)			84BL	30
N002802	1.047	(0.098)	-2.334	(0.169)	0.297	(0.057)			84BL	31
N002804	0.280	(0.043)	1.634	(0.273)	0.000	(0.000)			84BL	32
N002902	0.731	(0.078)	-1.329	(0.205)	0.375	(0.057)			84BM	6
N002903	1.283	(0.119)	-1.481	(0.111)	0.290	(0.043)			84BM	7
N002904	0.790	(0.077)	-1.144	(0.162)	0.333	(0.050)			84BM	8
N002905	0.520	(0.064)	-0.168	(0.190)	0.279	(0.048)			84BM	9
N002906	1.308	(0.118)	-1.382	(0.102)	0.279	(0.042)			84BM	10
N003001	1.010	(0.101)	0.409	(0.069)	0.214	(0.028)			84BM	11
N003002	0.362	(0.050)	-0.198	(0.274)	0.227	(0.055)			84BM	12
N003003	1.854	(0.146)	1.102	(0.037)	0.089	(0.010)			84BM	13
N003101	0.744	(0.074)	-2.219	(0.233)	0.334	(0.064)			84BM	14
N003102	1.154	(0.106)	-1.249	(0.115)	0.323	(0.044)			84BM	15
N003104	0.653	(0.046)	1.019	(0.069)	0.000	(0.000)			84BM	16
N003201	1.053	(0.104)	-1.666	(0.154)	0.344	(0.052)			84BN	21
N003202	0.822	(0.080)	-1.059	(0.154)	0.348	(0.049)			84BN	22
N003203	1.192	(0.094)	-0.372	(0.071)	0.253	(0.032)			84BN	23
N003204	0.939	(0.079)	-1.103	(0.115)	0.263	(0.043)			84BN	24
N003301	0.794	(0.071)	-1.572	(0.162)	0.277	(0.051)			84BN	25
N003501	0.625	(0.065)	-1.058	(0.197)	0.319	(0.053)			84BN	27
N003601	1.094	(0.109)	-1.565	(0.146)	0.355	(0.051)			84BN	28
N003602	1.121	(0.096)	-0.926	(0.100)	0.308	(0.041)			84BN	29
N003701	0.738	(0.072)	-1.236	(0.175)	0.314	(0.053)			84BN	30
N003702	1.785	(0.159)	-0.377	(0.061)	0.369	(0.032)			84BN	31
N003704	0.704	(0.063)	-0.862	(0.095)	0.000	(0.000)			84BN	32
N003801	0.774	(0.103)	0.712	(0.100)	0.279	(0.032)			84BO	12
N003802	0.221	(0.033)	-1.941	(0.513)	0.229	(0.061)			84BO	13
N003803	0.712	(0.156)	1.718	(0.177)	0.308	(0.027)			84BO	14
N004201	1.142	(0.107)	-0.375	(0.091)	0.359	(0.036)			84BO	21
N004202	0.667	(0.079)	-0.192	(0.159)	0.335	(0.045)			84BO	22

BEST COPY AVAILABLE

Table E-3 (continued)
IRT Parameters for Reading Long-term Trend Samples
Age 17

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
N004301	0.986	(0.095)	-0.450	(0.106)	0.313	(0.040)			84BO	23
N004303	0.541	(0.045)	0.012	(0.069)	0.000	(0.000)			84BO	24
N004501	0.686	(0.077)	-0.331	(0.154)	0.319	(0.045)			84BP	20
N004502	0.541	(0.054)	-1.729	(0.236)	0.286	(0.057)			84BP	21
N004601	0.795	(0.074)	-0.197	(0.103)	0.236	(0.036)			84BP	22
N004602	1.106	(0.091)	-0.800	(0.089)	0.260	(0.037)			84BP	23
N004603	1.265	(0.114)	-1.076	(0.098)	0.315	(0.041)			84BP	24
N004605	0.538	(0.042)	-1.322	(0.113)	0.000	(0.000)			84BP	25
N004901	0.715	(0.065)	-1.179	(0.156)	0.268	(0.047)			84BQ	10
N005001	1.867	(0.172)	0.649	(0.036)	0.221	(0.017)			84BQ	7
N005002	1.068	(0.123)	0.678	(0.069)	0.288	(0.025)			84BQ	8
N005003	0.689	(0.094)	1.368	(0.105)	0.121	(0.023)			84BQ	9
N005201	1.252	(0.225)	0.629	(0.107)	0.654	(0.023)			84BQ	11
N005202	0.497	(0.071)	0.067	(0.214)	0.325	(0.050)			84BQ	12
N005203	0.569	(0.098)	1.109	(0.150)	0.287	(0.036)			84BQ	13
N005503	0.727	(0.081)	-0.103	(0.133)	0.322	(0.040)			84BR	14
N005504	1.494	(0.140)	0.222	(0.055)	0.323	(0.025)			84BR	15
N005505	0.872	(0.096)	-1.839	(0.223)	0.413	(0.062)			84BR	16
N015101	1.081	(0.119)	0.228	(0.085)	0.393	(0.030)			84BR	17
N015102	2.915	(0.231)	-0.022	(0.028)	0.224	(0.019)			84BR	18
N015103	2.558	(0.195)	0.050	(0.029)	0.203	(0.018)			84BR	19
N015104	1.749	(0.147)	-0.176	(0.053)	0.326	(0.027)			84BR	20
N015201	0.683	(0.065)	-2.249	(0.229)	0.311	(0.062)			84BN	26
N015502	1.319	(0.122)	-0.234	(0.077)	0.375	(0.032)			84BP	16
N015503	1.286	(0.143)	0.346	(0.071)	0.387	(0.027)			84BP	17
N015504	1.232	(0.105)	-0.453	(0.079)	0.319	(0.035)			84BP	18
N015505	0.804	(0.098)	-0.221	(0.150)	0.434	(0.042)			84BP	19
N015901	1.037	(0.125)	0.041	(0.111)	0.473	(0.034)			84BQ	14
N015902	1.460	(0.139)	0.092	(0.062)	0.355	(0.027)			84BQ	15
N015903	1.421	(0.137)	0.501	(0.050)	0.240	(0.022)			84BQ	16
N016001	0.833	(0.075)	-0.816	(0.126)	0.288	(0.043)			84BO	15
N016002	1.421	(0.152)	0.452	(0.059)	0.353	(0.024)			84BO	16
N016003	0.864	(0.079)	-0.237	(0.098)	0.270	(0.035)			84BO	17
N016004	0.966	(0.081)	-0.617	(0.096)	0.265	(0.037)			84BO	18
N016005	1.664	(0.129)	-0.509	(0.055)	0.270	(0.030)			84BO	19
N016006	0.883	(0.077)	-0.277	(0.092)	0.235	(0.035)			84BO	20
N017001	1.202	(0.106)	-0.180	(0.075)	0.323	(0.031)			84BH	7
N017002	1.565	(0.138)	0.465	(0.043)	0.222	(0.020)			84BH	8
N017003	1.961	(0.156)	1.169	(0.041)	0.192	(0.012)			84BH	9

Table E-4
IRT Parameters for Reading Main Samples
Reading for Literary Experience

									9/4		8/13		12/17	
NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item	Block	Item	Block	Item
R012001	2.005	(0.148)	-0.097	(0.025)	0.103	(0.013)			92R3	1	--	--	--	--
R012002	1.873	(0.084)	-0.710	(0.019)	0.000	(0.000)			92R3	2	--	--	--	--
R012003	1.720	(0.111)	-1.140	(0.043)	0.151	(0.026)			92R3	3	--	--	--	--
R012004	0.962	(0.055)	-0.201	(0.034)	0.000	(0.000)			92R3	4	--	--	--	--
R012005	1.323	(0.103)	-0.548	(0.045)	0.168	(0.024)			92R3	5	--	--	--	--
R012006	0.623	(0.034)	0.079	(0.032)			0.442	(0.059)	92R3	6	--	--	--	--
							0.024	(0.070)						
							-0.467	(0.092)						
R012007	1.078	(0.097)	-1.019	(0.087)	0.244	(0.039)			92R3	7	--	--	--	--
R012008	1.167	(0.062)	-1.099	(0.036)	0.000	(0.000)			92R3	8	--	--	--	--
R012009	1.898	(0.170)	-1.159	(0.062)	0.317	(0.036)			92R3	9	--	--	--	--
R012010	1.402	(0.075)	-1.060	(0.032)	0.000	(0.000)			92R3	10	--	--	--	--
R012011	2.240	(0.198)	-0.779	(0.042)	0.253	(0.028)			92R3	11	--	--	--	--
R012101	2.112	(0.144)	-1.392	(0.051)	0.326	(0.033)			92R4	1	--	--	--	--
R012102	0.998	(0.051)	-1.577	(0.046)	0.000	(0.000)			92R4	2	--	--	--	--
R012103	1.468	(0.099)	-1.104	(0.051)	0.165	(0.028)			92R4	3	--	--	--	--
R012104	0.895	(0.048)	-0.887	(0.036)	0.000	(0.000)			92R4	4	--	--	--	--
R012105	0.959	(0.084)	-0.663	(0.071)	0.166	(0.031)			92R4	5	--	--	--	--
R012106	1.029	(0.057)	-0.267	(0.032)	0.000	(0.000)			92R4	6	--	--	--	--
R012107	1.510	(0.133)	-0.508	(0.048)	0.236	(0.026)			92R4	7	--	--	--	--
R012108	0.772	(0.047)	-1.638	(0.064)	0.000	(0.000)			92R4	8	--	--	--	--
R012109	0.773	(0.047)	-1.410	(0.057)	0.000	(0.000)			92R4	9	--	--	--	--
R012110	0.942	(0.085)	-1.611	(0.127)	0.243	(0.051)			92R4	10	--	--	--	--
R012111	1.121	(0.059)	0.707	(0.034)			0.890	(0.033)	92R4	11	--	--	--	--
							-0.890	(0.101)						
R012112	0.923	(0.061)	-1.128	(0.051)	0.000	(0.000)			92R4	12	--	--	--	--
R012401	1.055	(0.049)	1.020	(0.032)			1.261	(0.031)	92R9	1	--	--	--	--
							-0.416	(0.083)						
							-0.845	(0.318)						
R012402	1.007	(0.106)	-0.760	(0.096)	0.297	(0.040)			92R9	2	--	--	--	--
R012403	1.314	(0.076)	0.127	(0.035)	0.000	(0.000)			92R9	3	--	--	--	--
R012404	1.224	(0.109)	-0.523	(0.056)	0.201	(0.027)			92R9	4	--	--	--	--
R012405	1.395	(0.157)	0.080	(0.047)	0.202	(0.020)			92R9	5	--	--	--	--
R012406	1.059	(0.060)	-0.317	(0.032)	0.000	(0.000)			92R9	6	--	--	--	--
R012407	1.171	(0.061)	-0.909	(0.032)	0.000	(0.000)			92R9	7	--	--	--	--
R012408	1.705	(0.165)	-0.478	(0.048)	0.277	(0.026)			92R9	8	--	--	--	--
R012409	1.560	(0.092)	-0.095	(0.028)	0.000	(0.000)			92R9	9	--	--	--	--
R012601	1.209	(0.042)	0.291	(0.021)	0.000	(0.000)			92R5	1	92R5	1	--	--
R012602	1.626	(0.096)	0.423	(0.026)	0.151	(0.011)			92R5	2	92R5	2	--	--
R012603	1.754	(0.087)	-0.521	(0.030)	0.177	(0.019)			92R5	3	92R5	3	--	--
R012604	1.117	(0.040)	0.383	(0.023)	0.000	(0.000)			92R5	4	92R5	4	--	--
R012605	1.059	(0.076)	0.077	(0.053)	0.259	(0.023)			92R5	5	92R5	5	--	--
R012606	1.891	(0.103)	-0.379	(0.030)	0.233	(0.019)			92R5	6	92R5	6	--	--
R012607	1.346	(0.066)	0.915	(0.023)			1.039	(0.029)	92R5	7	--	--	--	--
							0.638	(0.040)						
							-1.678	(0.384)						
R012607	0.875	(0.033)	0.992	(0.023)			1.130	(0.044)	--	--	92R5	7	--	--
							0.620	(0.040)						
							-1.750	(0.110)						
R012608	0.735	(0.050)	-1.200	(0.125)	0.235	(0.048)			92R5	8	92R5	8	--	--
R012609	1.692	(0.128)	0.018	(0.047)	0.213	(0.029)			--	--	92R5	9	--	--
R012609	0.930	(0.113)	-0.101	(0.074)	0.184	(0.031)			92R5	9	--	--	--	--
R012610	1.929	(0.133)	-0.027	(0.035)	0.369	(0.018)			92R5	10	92R5	10	--	--
R012611	0.992	(0.040)	-0.903	(0.032)	0.000	(0.000)			92R5	11	92R5	11	--	--

Table E-4 (continued)
IRT Parameters for Reading Main Samples
Reading for Literary Experience

									9/4		8/13		12/17	
NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item	Block	Item	Block	Item
R012801	0.840	(0.073)	0.051	(0.088)	0.171	(0.038)			--	--	92R4	1	--	--
R012802	1.731	(0.123)	-0.673	(0.056)	0.201	(0.038)			--	--	92R4	2	--	--
R012803	1.108	(0.057)	0.613	(0.032)	0.000	(0.000)			--	--	92R4	3	--	--
R012804	1.061	(0.097)	-0.206	(0.095)	0.263	(0.045)			--	--	92R4	4	--	--
R012805	1.093	(0.098)	0.297	(0.063)	0.187	(0.030)			--	--	92R4	5	--	--
R012806	1.270	(0.097)	-0.262	(0.067)	0.214	(0.036)			--	--	92R4	6	--	--
R012807	0.867	(0.051)	0.787	(0.044)	0.000	(0.000)			--	--	92R4	7	--	--
R012808	0.940	(0.049)	1.473	(0.038)			0.350	(0.041)	--	--	92R4	8	--	--
							0.175	(0.065)						
							-0.525	(0.119)						
R012809	1.129	(0.093)	-0.416	(0.088)	0.243	(0.045)			--	--	92R4	9	--	--
R012810	0.953	(0.050)	0.051	(0.032)	0.000	(0.000)			--	--	92R4	10	--	--
R012811	1.299	(0.105)	-0.364	(0.075)	0.235	(0.042)			--	--	92R4	11	--	--
R012812	1.090	(0.114)	0.196	(0.084)	0.291	(0.037)			--	--	92R4	12	--	--
R012813	0.730	(0.051)	0.656	(0.051)	0.000	(0.000)			--	--	92R4	13	--	--
R013101	0.928	(0.049)	0.423	(0.046)	0.103	(0.021)			--	--	92R3	1	92R3	1
R013102	1.024	(0.037)	0.832	(0.023)	0.000	(0.000)			--	--	92R3	2	92R3	2
R013103	0.912	(0.059)	-0.298	(0.096)	0.263	(0.043)			--	--	92R3	3	92R3	3
R013104	1.252	(0.041)	-0.031	(0.021)	0.000	(0.000)			--	--	92R3	4	92R3	4
R013105	0.931	(0.058)	1.084	(0.052)	0.000	(0.000)			--	--	92R3	5	--	--
R013105	0.919	(0.047)	0.546	(0.034)	0.000	(0.000)			--	--	--	--	92R3	5
R013106	1.125	(0.040)	1.386	(0.023)			1.913	(0.044)	--	--	--	--	92R3	6
							0.152	(0.031)						
							-2.065	(0.136)						
R013106	1.251	(0.045)	1.414	(0.024)			1.856	(0.032)	--	--	92R3	6	--	--
							-0.040	(0.039)						
							-1.816	(0.290)						
R013107	1.364	(0.087)	0.348	(0.048)	0.289	(0.024)			--	--	92R3	7	92R3	7
R013108	0.819	(0.071)	3.061	(0.178)	0.000	(0.000)			--	--	92R3	8	92R3	8
R013109	1.267	(0.074)	0.818	(0.032)	0.114	(0.015)			--	--	92R3	9	92R3	9
R013110	0.528	(0.029)	0.725	(0.042)	0.000	(0.000)			--	--	92R3	10	92R3	10
R013111	1.085	(0.049)	1.365	(0.033)	0.000	(0.000)			--	--	92R3	11	92R3	11
R013501	0.973	(0.049)	0.232	(0.037)	0.000	(0.000)			--	--	--	--	92R4	1
R013502	1.421	(0.110)	0.140	(0.070)	0.264	(0.038)			--	--	--	--	92R4	2
R013503	0.567	(0.039)	-0.334	(0.080)	0.000	(0.000)			--	--	--	--	92R4	3
R013504	0.746	(0.068)	0.500	(0.107)	0.191	(0.040)			--	--	--	--	92R4	4
R013505	0.897	(0.052)	-0.363	(0.058)	0.000	(0.000)			--	--	--	--	92R4	5
R013506	0.419	(0.022)	1.785	(0.068)			-1.745	(0.141)	--	--	--	--	92R4	6
							1.745	(0.158)						
R013507	1.037	(0.096)	0.827	(0.069)	0.174	(0.031)			--	--	--	--	92R4	7
R013508	0.474	(0.046)	1.339	(0.089)	0.000	(0.000)			--	--	--	--	92R4	8
R013509	0.937	(0.063)	0.780	(0.043)	0.000	(0.000)			--	--	--	--	92R4	9
R013601	1.492	(0.129)	0.681	(0.056)	0.262	(0.029)			--	--	--	--	92R5	1
R013602	0.797	(0.044)	0.854	(0.037)	0.000	(0.000)			--	--	--	--	92R5	2
R013603	1.333	(0.104)	0.530	(0.060)	0.212	(0.031)			--	--	--	--	92R5	3
R013604	0.571	(0.039)	-0.482	(0.085)	0.000	(0.000)			--	--	--	--	92R5	4
R013605	0.921	(0.047)	0.681	(0.033)	0.000	(0.000)			--	--	--	--	92R5	5
R013606	0.688	(0.092)	1.242	(0.117)	0.237	(0.040)			--	--	--	--	92R5	6
R013607	0.780	(0.044)	0.092	(0.047)	0.000	(0.000)			--	--	--	--	92R5	7
R013608	1.090	(0.100)	1.245	(0.062)	0.277	(0.024)			--	--	--	--	92R5	8
R013609	1.053	(0.055)	-0.063	(0.041)	0.000	(0.000)			--	--	--	--	92R5	9
R013610	1.078	(0.060)	1.309	(0.035)	0.000	(0.000)			--	--	--	--	92R5	10
R013611	0.796	(0.049)	1.252	(0.045)	0.000	(0.000)			--	--	--	--	92R5	11
R013612	1.027	(0.112)	1.059	(0.073)	0.216	(0.032)			--	--	--	--	92R5	12

Table E-5
IRT Parameters for Reading Main Samples
Reading to Gain Information

									9/4		8/13		12/17	
NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item	Block	Item	Block	Item
R012201	0.452	(0.035)	-0.874	(0.060)	0.000	(0.000)			92R6	1	--	--	--	--
R012202	0.942	(0.105)	-0.509	(0.076)	0.240	(0.031)			92R6	2	--	--	--	--
R012203	0.872	(0.095)	-0.406	(0.073)	0.178	(0.029)			92R6	3	--	--	--	--
R012204	0.535	(0.025)	-0.651	(0.031)			1.153	(0.069)	92R6	4	--	--	--	--
							-0.283	(0.066)						
							-0.870	(0.086)						
R012205	1.659	(0.162)	-0.484	(0.044)	0.260	(0.022)			92R6	5	--	--	--	--
R012206	1.406	(0.077)	-0.266	(0.030)	0.000	(0.000)			92R6	6	--	--	--	--
R012207	0.727	(0.073)	-1.364	(0.130)	0.225	(0.046)			92R6	7	--	--	--	--
R012208	0.921	(0.054)	-1.134	(0.039)	0.000	(0.000)			92R6	8	--	--	--	--
R012209	2.009	(0.203)	-0.537	(0.041)	0.268	(0.022)			92R6	9	--	--	--	--
R012210	0.797	(0.054)	-2.018	(0.074)	0.000	(0.000)			92R6	10	--	--	--	--
R012301	0.895	(0.094)	-1.024	(0.106)	0.292	(0.042)			92R8	1	--	--	--	--
R012302	1.211	(0.090)	-1.284	(0.063)	0.204	(0.033)			92R8	2	--	--	--	--
R012303	1.328	(0.062)	-1.229	(0.026)	0.000	(0.000)			92R8	3	--	--	--	--
R012304	2.409	(0.669)	0.816	(0.157)	0.248	(0.012)			92R8	4	--	--	--	--
R012305	0.563	(0.020)	-0.041	(0.038)			2.408	(0.071)	92R8	5	--	--	--	--
							0.048	(0.058)						
							-2.457	(0.177)						
R012306	0.939	(0.057)	-0.169	(0.044)	0.000	(0.000)			92R8	6	--	--	--	--
R012307	1.592	(0.114)	-1.003	(0.043)	0.172	(0.025)			92R8	7	--	--	--	--
R012308	1.036	(0.058)	-0.526	(0.033)	0.000	(0.000)			92R8	8	--	--	--	--
R012309	1.074	(0.152)	-0.127	(0.072)	0.254	(0.028)			92R8	9	--	--	--	--
R012310	1.046	(0.066)	-0.460	(0.037)	0.000	(0.000)			92R8	10	--	--	--	--
R012501	0.623	(0.197)	1.439	(0.433)	0.269	(0.026)			92R10	1	--	--	--	--
R012502	1.163	(0.085)	-2.325	(0.105)	0.238	(0.049)			92R10	2	--	--	--	--
R012503	1.292	(0.065)	-0.450	(0.027)	0.000	(0.000)			92R10	3	--	--	--	--
R012504	0.990	(0.051)	-0.964	(0.031)	0.000	(0.000)			92R10	4	--	--	--	--
R012505	1.416	(0.095)	-1.471	(0.059)	0.222	(0.033)			92R10	5	--	--	--	--
R012506	1.073	(0.055)	-0.988	(0.030)	0.000	(0.000)			92R10	6	--	--	--	--
R012507	1.491	(0.113)	-1.257	(0.056)	0.253	(0.032)			92R10	7	--	--	--	--
R012508	1.268	(0.063)	-1.114	(0.028)	0.000	(0.000)			92R10	8	--	--	--	--
R012509	0.784	(0.077)	-1.550	(0.135)	0.250	(0.050)			92R10	9	--	--	--	--
R012510	1.068	(0.107)	-1.099	(0.092)	0.309	(0.040)			92R10	10	--	--	--	--
R012511	1.379	(0.073)	-1.247	(0.030)	0.000	(0.000)			92R10	11	--	--	--	--
R012512	0.586	(0.032)	-0.165	(0.036)			0.867	(0.066)	92R10	12	--	--	--	--
							0.220	(0.071)						
							-1.087	(0.113)						
R012701	1.516	(0.085)	-0.935	(0.047)	0.297	(0.025)			92R7	1	92R7	1	--	--
R012702	1.038	(0.059)	-1.350	(0.060)	0.000	(0.000)			--	--	92R7	2	--	--
R012702	0.792	(0.045)	-1.771	(0.053)	0.000	(0.000)			92R7	2	--	--	--	--
R012703	1.176	(0.039)	-0.196	(0.020)	0.000	(0.000)			92R7	3	92R7	8	--	--
R012704	2.074	(0.108)	-0.254	(0.023)	0.187	(0.013)			92R7	4	92R7	4	--	--
R012705	1.436	(0.122)	0.606	(0.077)	0.000	(0.000)			92R7	5	--	--	--	--
R012705	1.294	(0.066)	0.412	(0.026)	0.000	(0.000)			--	--	92R7	5	--	--
R012706	0.722	(0.030)	0.024	(0.032)	0.000	(0.000)			92R7	6	92R7	6	--	--
R012707	2.385	(0.138)	-0.551	(0.027)	0.283	(0.017)			92R7	7	92R7	3	--	--
R012708	0.773	(0.027)	0.937	(0.029)			1.707	(0.045)	--	--	92R7	10	--	--
							-0.276	(0.050)						
							-1.430	(0.138)						
R012708	1.168	(0.057)	0.657	(0.030)			1.253	(0.032)	92R7	8	--	--	--	--
							-0.115	(0.063)						
							-1.138	(0.298)						
R012709	0.746	(0.057)	-0.915	(0.120)	0.261	(0.044)			92R7	9	92R7	9	--	--

BEST COPY AVAILABLE

Table E-5 (continued)
IRT Parameters for Reading Main Samples
Reading to Gain Information

									9/4		8/13		12/17	
NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item	Block	Item	Block	Item
R012710	1.452	(0.051)	-0.436	(0.020)	0.000	(0.000)			92R7	10	92R7	11	--	--
R012711	1.619	(0.110)	-0.213	(0.042)	0.169	(0.026)			--	--	92R7	7	--	--
R012712	1.282	(0.130)	0.186	(0.062)	0.292	(0.030)			--	--	92R7	12	--	--
R012713	1.772	(0.090)	-0.351	(0.026)	0.000	(0.000)			--	--	92R7	13	--	--
R012901	1.016	(0.052)	0.026	(0.030)	0.000	(0.000)			--	--	92R8	1	--	--
R012902	0.524	(0.057)	-0.535	(0.185)	0.229	(0.053)			--	--	92R8	2	--	--
R012903	0.778	(0.034)	0.654	(0.025)			0.970	(0.044)	--	--	92R8	3	--	--
							-0.161	(0.049)						
							-0.808	(0.079)						
R012904	1.506	(0.176)	0.676	(0.045)	0.262	(0.021)			--	--	92R8	4	--	--
R012905	1.006	(0.055)	0.283	(0.031)	0.000	(0.000)			--	--	92R8	5	--	--
R012906	2.060	(0.143)	0.019	(0.031)	0.177	(0.020)			--	--	92R8	6	--	--
R012907	0.867	(0.050)	0.133	(0.034)	0.000	(0.000)			--	--	92R8	7	--	--
R012908	1.590	(0.132)	-0.195	(0.056)	0.283	(0.031)			--	--	92R8	8	--	--
R012909	1.196	(0.060)	-0.553	(0.034)	0.000	(0.000)			--	--	92R8	9	--	--
R012910	0.977	(0.085)	1.488	(0.098)	0.000	(0.000)			--	--	92R8	10	--	--
R012911	1.937	(0.155)	0.363	(0.031)	0.169	(0.018)			--	--	92R8	11	--	--
R012912	1.759	(0.139)	-0.647	(0.059)	0.250	(0.036)			--	--	92R8	12	--	--
R012913	1.445	(0.127)	-0.408	(0.072)	0.293	(0.038)			--	--	92R8	13	--	--
R012914	0.872	(0.059)	0.419	(0.040)	0.000	(0.000)			--	--	92R8	14	--	--
R013201	0.791	(0.024)	0.863	(0.016)			0.899	(0.031)	--	--	92R6	1	92R6	1
							-0.002	(0.032)						
							-0.897	(0.049)						
R013202	1.016	(0.068)	-0.088	(0.074)	0.265	(0.034)			--	--	92R6	2	92R6	2
R013203	1.172	(0.049)	-1.191	(0.045)	0.000	(0.000)			--	--	92R6	3	92R6	3
R013204	1.343	(0.087)	-0.077	(0.059)	0.369	(0.029)			--	--	92R6	4	92R6	4
R013205	0.772	(0.032)	-0.338	(0.036)	0.000	(0.000)			--	--	92R6	5	92R6	5
R013206	0.930	(0.056)	0.152	(0.057)	0.152	(0.027)			--	--	92R6	6	92R6	6
R013207	1.099	(0.040)	-0.253	(0.026)	0.000	(0.000)			--	--	92R6	7	92R6	7
R013208	2.502	(0.134)	-0.047	(0.026)	0.252	(0.021)			--	--	92R6	8	92R6	8
R013209	0.886	(0.040)	0.984	(0.030)	0.000	(0.000)			--	--	92R6	9	92R6	9
R013210	1.102	(0.126)	1.768	(0.089)	0.245	(0.013)			--	--	92R6	10	92R6	10
R013211	1.065	(0.042)	0.088	(0.024)	0.000	(0.000)			--	--	92R6	11	92R6	11
R013212	0.779	(0.021)	1.635	(0.028)			1.980	(0.036)	--	--	92R6	12	92R6	12
							-0.793	(0.059)						
							-1.187	(0.180)						
R013701	1.418	(0.084)	-0.367	(0.047)	0.000	(0.000)			--	--	--	--	92R7	1
R013702	1.720	(0.085)	0.127	(0.026)	0.000	(0.000)			--	--	--	--	92R7	2
R013703	1.267	(0.115)	-0.477	(0.111)	0.320	(0.055)			--	--	--	--	92R7	3
R013704	1.227	(0.075)	1.320	(0.037)	0.000	(0.000)			--	--	--	--	92R7	4
R013705	1.568	(0.120)	-0.183	(0.069)	0.246	(0.045)			--	--	--	--	92R7	5
R013706	0.747	(0.053)	0.221	(0.047)	0.000	(0.000)			--	--	--	--	92R7	6
R013707	1.202	(0.102)	0.680	(0.052)	0.155	(0.030)			--	--	--	--	92R7	7
R013708	0.713	(0.054)	-0.220	(0.072)	0.000	(0.000)			--	--	--	--	92R7	8
R013709	0.566	(0.111)	1.482	(0.172)	0.256	(0.049)			--	--	--	--	92R7	9
R013710	0.724	(0.058)	1.019	(0.046)	0.000	(0.000)			--	--	--	--	92R7	10
R013711	0.760	(0.161)	1.622	(0.135)	0.279	(0.041)			--	--	--	--	92R7	11
R013712	1.314	(0.083)	0.181	(0.037)	0.000	(0.000)			--	--	--	--	92R7	12
R013801	1.359	(0.070)	0.421	(0.025)	0.000	(0.000)			--	--	--	--	92R8	1
R013802	1.223	(0.107)	1.253	(0.057)	0.275	(0.023)			--	--	--	--	92R8	2
R013803	1.710	(0.079)	0.946	(0.019)	0.000	(0.000)			--	--	--	--	92R8	3
R013804	0.738	(0.071)	0.028	(0.119)	0.205	(0.045)			--	--	--	--	92R8	4
R013805	0.783	(0.041)	1.401	(0.037)			-0.423	(0.055)	--	--	--	--	92R8	5
							0.423	(0.067)						

Table E-5 (continued)
IRT Parameters for Reading Main Samples
Reading to Gain Information

									9/4		8/13		12/17	
NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item	Block	Item	Block	Item
R013806	0.979	(0.060)	0.819	(0.030)	0.000	(0.000)			–	–	–	–	92R8	6
R013807	1.505	(0.100)	1.025	(0.038)	0.216	(0.021)			–	–	–	–	92R8	7
R013808	1.229	(0.072)	1.012	(0.027)	0.000	(0.000)			–	–	–	–	92R8	8
R013809	1.103	(0.068)	0.948	(0.029)	0.000	(0.000)			–	–	–	–	92R8	9
R013810	0.880	(0.059)	0.253	(0.043)	0.000	(0.000)			–	–	–	–	92R8	10
R014701	0.457	(0.097)	1.017	(0.233)	0.342	(0.050)			–	–	92R13	1	–	–
R014702	1.054	(0.053)	-0.214	(0.031)	0.000	(0.000)			–	–	92R13	2	–	–
R014703	1.119	(0.075)	1.083	(0.054)	0.000	(0.000)			–	–	92R13	3	–	–
R014704	1.398	(0.066)	0.068	(0.023)	0.000	(0.000)			–	–	92R13	4	–	–
R014705	0.629	(0.031)	0.347	(0.025)			0.587	(0.060)	–	–	92R13	5	–	–
							0.094	(0.059)						
							-0.681	(0.071)						
R014706	1.296	(0.061)	-0.362	(0.028)	0.000	(0.000)			–	–	92R13	6	–	–
R014707	1.413	(0.156)	0.714	(0.044)	0.192	(0.020)			–	–	92R13	7	–	–
R014708	0.828	(0.072)	-0.588	(0.112)	0.215	(0.045)			–	–	92R13	8	–	–
R014709	1.213	(0.059)	-0.458	(0.032)	0.000	(0.000)			–	–	92R13	9	–	–
R014710	0.860	(0.059)	0.927	(0.057)	0.000	(0.000)			–	–	92R13	10	–	–
R014711	0.891	(0.092)	-0.371	(0.122)	0.283	(0.049)			–	–	92R13	11	–	–
R014712	1.301	(0.292)	1.347	(0.119)	0.290	(0.019)			–	–	92R13	12	–	–
R014713	0.740	(0.029)	0.631	(0.029)			1.342	(0.049)	–	–	92R13	13	–	–
							-0.693	(0.058)						
							-0.649	(0.092)						
R015501	1.209	(0.104)	0.134	(0.082)	0.271	(0.043)			–	–	–	–	92R13	1
R015502	1.208	(0.112)	0.761	(0.056)	0.195	(0.031)			–	–	–	–	92R13	2
R015503	0.632	(0.051)	-1.286	(0.144)	0.000	(0.000)			–	–	–	–	92R13	3
R015504	1.192	(0.112)	0.572	(0.068)	0.251	(0.036)			–	–	–	–	92R13	4
R015505	0.843	(0.054)	-0.105	(0.054)	0.000	(0.000)			–	–	–	–	92R13	5
R015506	1.285	(0.122)	0.656	(0.061)	0.253	(0.034)			–	–	–	–	92R13	6
R015507	1.359	(0.049)	0.821	(0.017)			1.163	(0.039)	–	–	–	–	92R13	7
							-0.187	(0.026)						
							-0.976	(0.044)						
R015508	1.702	(0.137)	0.299	(0.054)	0.258	(0.036)			–	–	–	–	92R13	8
R015509	1.927	(0.092)	0.641	(0.018)	0.000	(0.000)			–	–	–	–	92R13	9
R015510	1.601	(0.132)	0.516	(0.049)	0.219	(0.032)			–	–	–	–	92R13	10
R015511	0.929	(0.090)	0.039	(0.113)	0.261	(0.049)			–	–	–	–	92R13	11
R015512	1.345	(0.073)	0.003	(0.035)	0.000	(0.000)			–	–	–	–	92R13	12
R015513	1.281	(0.106)	0.129	(0.075)	0.257	(0.041)			–	–	–	–	92R13	13
R015514	0.737	(0.031)	0.625	(0.026)			1.394	(0.072)	–	–	–	–	92R13	14
							-0.311	(0.047)						
							-1.084	(0.064)						
R015515	1.167	(0.095)	1.303	(0.049)	0.114	(0.019)			–	–	–	–	92R13	15
R015516	1.149	(0.104)	0.505	(0.068)	0.219	(0.036)			–	–	–	–	92R13	16
R015601	0.675	(0.028)	-0.247	(0.027)			2.412	(0.146)	–	–	–	–	92R14	1
							-0.868	(0.047)						
							-1.544	(0.053)						
R015602	1.248	(0.135)	0.594	(0.078)	0.351	(0.038)			–	–	–	–	92R14	2
R015603	0.813	(0.124)	0.907	(0.126)	0.373	(0.044)			–	–	–	–	92R14	3
R015604	1.738	(0.082)	0.277	(0.023)	0.000	(0.000)			–	–	–	–	92R14	4
R015605	1.997	(0.107)	0.996	(0.030)	0.278	(0.019)			–	–	–	–	92R14	5
R015606	0.922	(0.124)	1.744	(0.106)	0.245	(0.023)			–	–	–	–	92R14	6
R015607	1.855	(0.087)	0.596	(0.018)	0.000	(0.000)			–	–	–	–	92R14	7
R015608	1.348	(0.102)	1.241	(0.048)	0.234	(0.020)			–	–	–	–	92R14	8
R015609	1.108	(0.042)	0.375	(0.016)			1.031	(0.054)	–	–	–	–	92R14	9
							-0.615	(0.035)						
							-0.415	(0.036)						

Table E-5 (continued)
IRT Parameters for Reading Main Samples
Reading to Gain Information

									9/4		8/13		12/17	
NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item	Block	Item	Block	Item
R015610	1.270	(0.096)	0.364	(0.056)	0.169	(0.032)			--	--	--	--	92R14	10
R015611	1.095	(0.098)	0.072	(0.094)	0.273	(0.045)			--	--	--	--	92R14	11
R015612	0.541	(0.028)	0.891	(0.028)			0.862	(0.069)	--	--	--	--	92R14	12
							-0.125	(0.065)						
							-0.737	(0.081)						

Table E-6
IRT Parameters for Reading Main Samples
Reading to Perform a Task

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	9/4		8/13		12/17	
									Block	Item	Block	Item	Block	Item
R013001	0.976	(0.049)	-1.095	(0.042)	0.000	(0.000)			–	–	92R11	1	–	–
R013002	1.669	(0.120)	-0.527	(0.043)	0.188	(0.023)			–	–	92R11	2	–	–
R013003	1.021	(0.050)	-0.681	(0.034)	0.000	(0.000)			–	–	92R11	3	–	–
R013004	0.431	(0.027)	0.698	(0.064)			0.254	(0.074)	–	–	92R11	4	–	–
							-0.254	(0.099)						
R013005	0.790	(0.043)	-1.230	(0.054)	0.000	(0.000)			–	–	92R11	5	–	–
R013006	1.016	(0.087)	-0.551	(0.081)	0.245	(0.033)			–	–	92R11	6	–	–
R013007	0.781	(0.044)	-1.273	(0.057)	0.000	(0.000)			–	–	92R11	7	–	–
R013008	0.929	(0.050)	-0.172	(0.035)	0.000	(0.000)			–	–	92R11	8	–	–
R013009	0.845	(0.046)	-0.931	(0.047)	0.000	(0.000)			–	–	92R11	9	–	–
R013010	0.845	(0.047)	-0.967	(0.048)	0.000	(0.000)			–	–	92R11	10	–	–
R013011	0.619	(0.042)	0.173	(0.057)	0.000	(0.000)			–	–	92R11	11	–	–
R013012	1.190	(0.115)	-0.218	(0.069)	0.272	(0.029)			–	–	92R11	12	–	–
R013301	0.687	(0.062)	-0.258	(0.141)	0.391	(0.039)			–	–	92R9	1	92R9	1
R013302	1.167	(0.047)	1.127	(0.030)	0.000	(0.000)			–	–	92R9	2	92R9	2
R013303	1.062	(0.070)	-0.003	(0.061)	0.297	(0.025)			–	–	92R9	3	92R9	3
R013304	0.487	(0.049)	2.062	(0.213)	0.000	(0.000)			–	–	92R9	4	–	–
R013304	0.728	(0.049)	1.307	(0.059)	0.000	(0.000)			–	–	–	–	92R9	4
R013305	1.511	(0.146)	-1.356	(0.114)	0.291	(0.057)			–	–	–	–	92R9	5
R013305	1.414	(0.133)	-1.427	(0.100)	0.425	(0.044)			–	–	92R9	5	–	–
R013306	0.773	(0.054)	-0.552	(0.108)	0.286	(0.038)			–	–	92R9	6	92R9	6
R013307	0.927	(0.034)	0.501	(0.025)	0.000	(0.000)			–	–	92R9	7	92R9	7
R013308	1.359	(0.097)	0.186	(0.051)	0.408	(0.020)			–	–	92R9	8	92R9	8
R013309	0.905	(0.082)	0.846	(0.056)	0.252	(0.020)			–	–	92R9	9	92R9	9
R013310	0.799	(0.030)	0.392	(0.028)	0.000	(0.000)			–	–	92R9	10	92R9	10
R013311	1.105	(0.082)	0.093	(0.066)	0.388	(0.024)			–	–	92R9	11	92R9	11
R013312	0.479	(0.021)	1.879	(0.058)			0.124	(0.056)	–	–	92R9	12	92R9	12
							0.261	(0.087)						
							-0.385	(0.142)						
R013401	0.885	(0.050)	0.057	(0.052)	0.146	(0.022)			–	–	92R10	1	92R10	1
R013402	0.940	(0.032)	-0.290	(0.025)	0.000	(0.000)			–	–	92R10	2	92R10	2
R013403	0.500	(0.012)	0.298	(0.023)			-2.113	(0.082)	–	–	92R10	3	92R10	3
							2.113	(0.082)						
R013404	1.041	(0.070)	-0.325	(0.074)	0.359	(0.029)			–	–	92R10	4	92R10	4
R013405	0.980	(0.034)	-0.495	(0.027)	0.000	(0.000)			–	–	92R10	5	92R10	5
R013406	0.683	(0.029)	0.605	(0.034)	0.000	(0.000)			–	–	92R10	6	92R10	6
R013407	0.625	(0.026)	-0.729	(0.043)	0.000	(0.000)			–	–	92R10	7	92R10	7
R013408	0.759	(0.059)	0.151	(0.085)	0.260	(0.029)			–	–	92R10	8	92R10	8
R013409	0.871	(0.032)	-0.625	(0.032)	0.000	(0.000)			–	–	92R10	9	92R10	9
R013410	0.724	(0.073)	-0.814	(0.143)	0.287	(0.046)			–	–	92R10	10	–	–
R013410	0.741	(0.076)	-1.009	(0.205)	0.344	(0.066)			–	–	–	–	92R10	10
R013411	0.676	(0.030)	0.541	(0.036)	0.000	(0.000)			–	–	92R10	11	92R10	11
R013412	0.451	(0.028)	-2.107	(0.126)	0.000	(0.000)			–	–	92R10	12	92R10	12
R013901	1.950	(0.161)	0.685	(0.040)	0.378	(0.021)			–	–	–	–	92R11	1
R013902	0.780	(0.045)	-0.635	(0.061)	0.000	(0.000)			–	–	–	–	92R11	2
R013903	1.028	(0.052)	0.536	(0.030)	0.000	(0.000)			–	–	–	–	92R11	3
R013904	0.911	(0.053)	-0.922	(0.065)	0.000	(0.000)			–	–	–	–	92R11	4
R013905	0.890	(0.195)	2.467	(0.260)	0.213	(0.016)			–	–	–	–	92R11	5
R013906	0.505	(0.040)	1.255	(0.079)	0.000	(0.000)			–	–	–	–	92R11	6
R013907	1.341	(0.126)	0.515	(0.062)	0.300	(0.029)			–	–	–	–	92R11	7
R013908	0.754	(0.045)	-0.733	(0.067)	0.000	(0.000)			–	–	–	–	92R11	8
R013909	0.716	(0.072)	0.093	(0.133)	0.251	(0.046)			–	–	–	–	92R11	9
R013910	1.090	(0.054)	0.381	(0.029)	0.000	(0.000)			–	–	–	–	92R11	10
R013911	0.950	(0.090)	-0.517	(0.139)	0.347	(0.054)			–	–	–	–	92R11	11

Table E-6 (continued)
IRT Parameters for Reading Main Samples
Reading to Perform a Task

									9/4		8/13		12/17	
NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item	Block	Item	Block	Item
R013912	0.576	(0.112)	1.453	(0.159)	0.293	(0.042)			--	--	--	--	92R11	12
R013913	0.548	(0.038)	0.140	(0.055)	0.000	(0.000)			--	--	--	--	92R11	13
R013914	0.420	(0.062)	0.133	(0.282)	0.316	(0.060)			--	--	--	--	92R11	14
R013915	0.453	(0.024)	1.393	(0.059)			-2.228	(0.164)	--	--	--	--	92R11	15
							2.228	(0.176)						

Table E-7
IRT Parameters for Mathematics Long-term Trend Samples
Age 9

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
N250301	0.917	(0.091)	1.195	(0.058)	0.324	(0.016)			86M2	20
N250601	0.900	(0.047)	-1.352	(0.094)	0.227	(0.043)			86M2	13
N250602	0.672	(0.036)	-1.307	(0.115)	0.195	(0.044)			86M2	14
N250603	0.937	(0.048)	0.060	(0.046)	0.141	(0.020)			86M2	15
N250701	0.659	(0.037)	-1.358	(0.119)	0.169	(0.044)			86M1	7
N250702	1.195	(0.070)	0.515	(0.034)	0.162	(0.014)			86M1	8
N250703	1.040	(0.050)	-0.459	(0.050)	0.118	(0.024)			86M1	9
N250901	0.537	(0.033)	-1.643	(0.169)	0.215	(0.053)			86M2	17
N250902	1.163	(0.069)	0.619	(0.033)	0.182	(0.014)			86M2	18
N250903	1.165	(0.053)	-0.067	(0.036)	0.129	(0.017)			86M2	19
N251401	0.674	(0.040)	-0.749	(0.109)	0.199	(0.040)			86M2	16
N252001	1.274	(0.100)	1.704	(0.055)	0.202	(0.008)			86M2	25
N252101	0.725	(0.095)	1.913	(0.108)	0.201	(0.016)			86M1	25
N257201	0.961	(0.056)	-0.623	(0.077)	0.264	(0.033)			86M1	11
N257801	0.703	(0.042)	-1.148	(0.126)	0.261	(0.046)			86M2	3
N258501	0.555	(0.076)	1.869	(0.116)	0.221	(0.022)			86M3	19
N261401	0.439	(0.031)	-0.577	(0.161)	0.206	(0.042)			86M2	12
N262201	0.690	(0.049)	-0.706	(0.133)	0.314	(0.043)			86M1	10
N262401	0.745	(0.060)	0.383	(0.081)	0.292	(0.026)			86M3	18
N262501	0.496	(0.052)	0.021	(0.193)	0.324	(0.046)			86M1	19
N263401	0.799	(0.047)	-1.349	(0.124)	0.296	(0.049)			86M2	4
N263402	0.836	(0.051)	-0.630	(0.095)	0.301	(0.036)			86M2	5
N265401	0.281	(0.079)	5.093	(1.026)	0.278	(0.023)			86M1	21
N266101	0.701	(0.094)	1.490	(0.092)	0.312	(0.021)			86M1	22
N267001	0.791	(0.044)	-1.573	(0.119)	0.264	(0.049)			86M3	16
N267601	1.351	(0.071)	-0.562	(0.049)	0.250	(0.026)			86M1	3
N267602	1.099	(0.051)	-0.126	(0.039)	0.128	(0.018)			86M1	18
N268201	1.371	(0.088)	0.669	(0.032)	0.224	(0.013)			86M1	24
N269001	0.552	(0.123)	3.869	(0.502)	0.088	(0.010)			86M2	26
N269101	0.580	(0.085)	1.853	(0.122)	0.229	(0.022)			86M1	23
N270001	0.583	(0.024)	-0.596	(0.039)	0.000	(0.000)			86M1	14
N270901	0.776	(0.043)	-2.755	(0.109)	0.000	(0.000)			86M1	1
N271101	0.776	(0.025)	0.070	(0.024)	0.000	(0.000)			86M2	24
N272101	0.823	(0.049)	-0.941	(0.106)	0.295	(0.041)			86M3	17
N272102	0.876	(0.051)	-0.234	(0.067)	0.189	(0.028)			86M1	15
N272301	0.830	(0.050)	-2.385	(0.135)	0.224	(0.055)			86M2	1
N272801	0.821	(0.045)	-1.932	(0.117)	0.223	(0.051)			86M3	15
N273501	0.615	(0.047)	-0.974	(0.182)	0.363	(0.052)			86M2	6
N275401	1.016	(0.031)	-0.739	(0.025)	0.000	(0.000)			86M2	7
N276001	0.925	(0.029)	-0.838	(0.028)	0.000	(0.000)			86M2	21
N276002	0.956	(0.033)	0.929	(0.027)	0.000	(0.000)			86M2	22
N276101	1.013	(0.035)	-0.960	(0.030)	0.000	(0.000)			86M1	12
N276601	1.152	(0.065)	-0.970	(0.076)	0.336	(0.036)			86M2	2
N276801	0.583	(0.042)	-3.762	(0.215)	0.000	(0.000)			86M1	4
N276802	0.568	(0.030)	-2.536	(0.114)	0.000	(0.000)			86M1	5
N276803	0.554	(0.023)	-0.118	(0.035)	0.000	(0.000)			86M1	6
N277401	0.845	(0.048)	-1.968	(0.111)	0.190	(0.047)			86M1	2
N277501	0.813	(0.026)	-0.489	(0.026)	0.000	(0.000)			86M2	8
N277601	0.842	(0.028)	-1.094	(0.034)	0.000	(0.000)			86M2	9
N277602	0.707	(0.024)	-0.133	(0.027)	0.000	(0.000)			86M2	10
N277603	0.784	(0.025)	-0.307	(0.026)	0.000	(0.000)			86M2	11
N284001	0.787	(0.028)	-0.790	(0.034)	0.000	(0.000)			86M1	16
N284002	0.760	(0.038)	1.830	(0.069)	0.000	(0.000)			86M1	17
N286101	0.896	(0.031)	-0.815	(0.031)	0.000	(0.000)			86M1	13
N286102	0.916	(0.028)	0.079	(0.021)	0.000	(0.000)			86M2	23

Table E-8
IRT Parameters for Mathematics Long-term Trend Samples
Age 13

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
N250201	0.638	(0.049)	-1.261	(0.178)	0.334	(0.055)			86M2	19
N250701	0.474	(0.040)	-3.997	(0.306)	0.109	(0.043)			86M2	14
N250702	0.854	(0.042)	-1.251	(0.073)	0.109	(0.031)			86M2	15
N250703	0.610	(0.040)	-2.712	(0.163)	0.118	(0.042)			86M2	16
N250901	0.393	(0.034)	-3.239	(0.294)	0.170	(0.052)			86M1	25
N250902	0.958	(0.049)	-0.716	(0.059)	0.116	(0.028)			86M1	26
N250903	0.765	(0.046)	-2.175	(0.118)	0.102	(0.038)			86M1	27
N252001	1.112	(0.088)	0.816	(0.046)	0.252	(0.017)			86M2	40
N252101	0.806	(0.079)	0.513	(0.089)	0.334	(0.029)			86M1	41
N252901	1.124	(0.056)	-0.046	(0.038)	0.097	(0.018)			86M1	32
N253701	0.274	(0.029)	-0.825	(0.272)	0.325	(0.040)			86M2	22
N254001	0.789	(0.047)	-0.757	(0.090)	0.185	(0.038)			86M3	28
N254601	0.996	(0.081)	-1.663	(0.149)	0.462	(0.055)			86M1	16
N254602	0.928	(0.081)	1.131	(0.057)	0.206	(0.017)			86M1	46
N255701	1.142	(0.081)	1.004	(0.040)	0.136	(0.013)			86M1	50
N256101	0.807	(0.033)	-1.312	(0.047)	0.000	(0.000)			86M2	17
N256501	1.261	(0.090)	0.471	(0.044)	0.298	(0.018)			86M3	30
N256801	1.316	(0.092)	0.538	(0.040)	0.286	(0.017)			86M3	32
N257601	1.181	(0.042)	-0.599	(0.025)	0.000	(0.000)			86M1	35
N258801	1.136	(0.101)	0.725	(0.055)	0.353	(0.019)			86M1	38
N258802	1.893	(0.116)	0.496	(0.026)	0.227	(0.013)			86M2	31
N258803	1.423	(0.103)	1.128	(0.035)	0.162	(0.010)			86M2	41
N260101	1.379	(0.083)	0.090	(0.039)	0.223	(0.020)			86M1	43
N261001	0.771	(0.063)	0.487	(0.077)	0.250	(0.026)			86M1	47
N261201	0.488	(0.077)	2.031	(0.154)	0.209	(0.027)			86M2	38
N261301	0.454	(0.038)	0.995	(0.101)	0.102	(0.025)			86M2	37
N261501	0.610	(0.043)	-1.026	(0.149)	0.225	(0.049)			86M2	34
N261801	0.639	(0.044)	-0.235	(0.100)	0.193	(0.034)			86M2	35
N262201	0.455	(0.038)	-1.702	(0.246)	0.324	(0.058)			86M2	18
N262401	0.893	(0.057)	-0.600	(0.086)	0.244	(0.037)			86M1	28
N262501	0.492	(0.042)	-1.039	(0.211)	0.313	(0.054)			86M1	33
N263101	0.657	(0.028)	-0.514	(0.037)	0.000	(0.000)			86M1	39
N263401	0.854	(0.059)	-2.227	(0.146)	0.276	(0.054)			86M2	12
N263402	0.606	(0.042)	-2.134	(0.177)	0.253	(0.055)			86M2	13
N263501	0.932	(0.048)	-0.013	(0.045)	0.092	(0.020)			86M2	30
N264701	1.490	(0.088)	0.456	(0.030)	0.195	(0.014)			86M2	33
N265201	0.766	(0.062)	-2.421	(0.194)	0.328	(0.061)			86M1	36
N265202	0.852	(0.069)	-0.226	(0.102)	0.363	(0.036)			86M1	30
N265901	0.828	(0.070)	0.782	(0.065)	0.232	(0.022)			86M1	40
N265902	0.727	(0.078)	1.047	(0.085)	0.305	(0.024)			86M3	31
N266101	0.931	(0.062)	-0.637	(0.091)	0.305	(0.038)			86M3	27
N266801	0.608	(0.044)	-1.345	(0.165)	0.270	(0.053)			86M1	31
N267201	1.114	(0.090)	-0.658	(0.101)	0.466	(0.038)			86M1	23
N269001	0.980	(0.061)	0.117	(0.053)	0.154	(0.024)			86M1	44
N269101	1.004	(0.065)	-0.166	(0.064)	0.225	(0.029)			86M2	26
N269201	0.961	(0.044)	1.656	(0.054)	0.000	(0.000)			86M2	44
N269901	0.678	(0.054)	-0.290	(0.125)	0.301	(0.040)			86M3	29
N270301	0.389	(0.030)	-2.143	(0.242)	0.143	(0.055)			86M2	20
N270302	1.302	(0.099)	1.795	(0.054)	0.052	(0.006)			86M2	21
N273901	1.800	(0.103)	0.095	(0.029)	0.212	(0.016)			86M1	37
N274801	1.316	(0.106)	0.333	(0.052)	0.397	(0.020)			86M1	29
N275001	0.840	(0.034)	0.850	(0.035)	0.000	(0.000)			86M1	42
N275301	0.349	(0.029)	-2.628	(0.282)	0.178	(0.054)			86M3	25
N276801	0.331	(0.046)	-6.495	(0.844)	0.000	(0.000)			86M1	17

BEST COPY AVAILABLE

Table E-8 (continued)
IRT Parameters for Mathematics Long-term Trend Samples
Age 13

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
N276802	0.410	(0.042)	-4.890	(0.447)	0.000	(0.000)			86M1	18
N276803	0.348	(0.025)	-2.545	(0.177)	0.000	(0.000)			86M1	19
N277401	0.551	(0.045)	-3.595	(0.262)	0.171	(0.052)			86M2	8
N277601	0.600	(0.041)	-3.256	(0.181)	0.000	(0.000)			86M1	20
N277602	0.631	(0.031)	-1.698	(0.075)	0.000	(0.000)			86M1	21
N277603	0.575	(0.033)	-2.335	(0.114)	0.000	(0.000)			86M1	22
N277901	0.689	(0.043)	-3.097	(0.148)	0.000	(0.000)			86M2	9
N277902	0.648	(0.043)	-3.317	(0.171)	0.000	(0.000)			86M2	10
N277903	0.792	(0.041)	-2.332	(0.090)	0.000	(0.000)			86M2	11
N278901	1.376	(0.079)	0.262	(0.034)	0.206	(0.016)			86M2	32
N278902	1.236	(0.106)	1.016	(0.045)	0.284	(0.014)			86M2	29
N278903	1.796	(0.126)	0.893	(0.029)	0.222	(0.011)			86M2	42
N278904	0.827	(0.095)	1.657	(0.082)	0.206	(0.016)			86M1	49
N281401	0.670	(0.093)	2.208	(0.137)	0.161	(0.016)			86M2	39
N281901	0.853	(0.069)	-2.727	(0.189)	0.223	(0.061)			86M1	15
N282201	1.376	(0.098)	0.559	(0.039)	0.305	(0.016)			86M2	28
N282202	0.896	(0.065)	-0.361	(0.093)	0.334	(0.035)			86M3	26
N283101	1.897	(0.108)	1.174	(0.029)	0.131	(0.008)			86M1	51
N285701	0.579	(0.049)	-0.147	(0.138)	0.214	(0.043)			86M2	27
N286201	0.958	(0.059)	-0.821	(0.084)	0.248	(0.038)			86M1	24
N286301	1.572	(0.095)	0.413	(0.030)	0.199	(0.015)			86M1	45
N286501	1.007	(0.080)	1.092	(0.048)	0.150	(0.015)			86M1	48
N286502	1.121	(0.077)	1.176	(0.040)	0.102	(0.011)			86M2	43
N286601	1.171	(0.039)	-0.305	(0.021)	0.000	(0.000)			86M2	23
N286602	1.046	(0.036)	-0.408	(0.024)	0.000	(0.000)			86M2	24
N286603	1.221	(0.042)	0.564	(0.022)	0.000	(0.000)			86M2	25

Table E-9
IRT Parameters for Mathematics Long-term Trend Samples
Age 17

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
N251101	1.235	(0.042)	0.965	(0.025)	0.000	(0.000)			86M1	49
N251701	1.030	(0.060)	-0.187	(0.059)	0.221	(0.026)			86M2	41
N253901	1.213	(0.067)	-0.376	(0.052)	0.246	(0.026)			86M1	39
N253902	1.007	(0.105)	0.859	(0.070)	0.465	(0.018)			86M1	40
N253903	1.238	(0.088)	0.543	(0.045)	0.324	(0.017)			86M1	41
N253904	1.641	(0.120)	0.571	(0.037)	0.385	(0.014)			86M1	42
N254001	0.893	(0.056)	-0.955	(0.105)	0.290	(0.045)			86M2	21
N254301	1.046	(0.075)	0.250	(0.061)	0.328	(0.023)			86M1	33
N254601	1.203	(0.088)	-2.374	(0.117)	0.244	(0.058)			86M2	15
N254602	1.386	(0.071)	-0.285	(0.040)	0.206	(0.022)			86M1	27
N255501	0.698	(0.053)	0.173	(0.096)	0.243	(0.032)			86M3	33
N255601	2.063	(0.114)	1.430	(0.036)	0.349	(0.008)			86M2	45
N255701	1.321	(0.059)	-0.935	(0.042)	0.120	(0.025)			86M1	32
N255801	0.848	(0.036)	1.559	(0.049)	0.000	(0.000)			86M2	49
N256001	0.902	(0.030)	-0.116	(0.024)	0.000	(0.000)			86M3	34
N256101	0.910	(0.043)	-2.034	(0.065)	0.000	(0.000)			86M1	15
N256801	1.097	(0.059)	-0.608	(0.060)	0.208	(0.030)			86M1	36
N257101	0.389	(0.073)	2.303	(0.206)	0.247	(0.034)			86M3	35
N258801	1.194	(0.073)	-0.471	(0.064)	0.304	(0.030)			86M2	38
N258802	1.712	(0.091)	-0.434	(0.036)	0.230	(0.022)			86M1	26
N258803	1.199	(0.073)	0.074	(0.048)	0.257	(0.022)			86M1	37
N258804	0.705	(0.049)	-2.553	(0.175)	0.242	(0.061)			86M1	18
N259001	1.011	(0.033)	-0.452	(0.024)	0.000	(0.000)			86M2	31
N259901	0.984	(0.056)	-0.447	(0.068)	0.234	(0.031)			86M1	28
N260101	1.403	(0.073)	-1.250	(0.054)	0.197	(0.034)			86M2	20
N260601	1.648	(0.070)	-1.512	(0.030)	0.000	(0.000)			86M1	16
N260801	1.344	(0.041)	0.083	(0.018)	0.000	(0.000)			86M2	43
N260901	2.077	(0.104)	-0.077	(0.024)	0.189	(0.014)			86M1	35
N261001	0.757	(0.045)	-0.695	(0.098)	0.205	(0.039)			86M2	40
N261201	0.553	(0.041)	-0.220	(0.135)	0.211	(0.041)			86M2	26
N261301	0.555	(0.042)	0.266	(0.108)	0.167	(0.034)			86M2	28
N261501	0.690	(0.043)	-2.125	(0.145)	0.207	(0.055)			86M2	24
N261601	0.836	(0.115)	1.700	(0.092)	0.372	(0.016)			86M2	27
N261801	0.590	(0.035)	-1.488	(0.137)	0.196	(0.048)			86M2	25
N262301	0.603	(0.045)	-1.250	(0.187)	0.306	(0.060)			86M2	17
N262401	0.999	(0.061)	-1.394	(0.102)	0.282	(0.049)			86M1	17
N262501	0.557	(0.041)	-1.359	(0.191)	0.342	(0.055)			86M2	35
N262502	1.253	(0.109)	1.349	(0.047)	0.276	(0.011)			86M2	36
N262601	0.761	(0.057)	0.438	(0.074)	0.220	(0.026)			86M1	38
N263001	0.635	(0.026)	0.813	(0.039)	0.000	(0.000)			86M1	43
N263101	0.662	(0.026)	-0.955	(0.042)	0.000	(0.000)			86M2	37
N263201	0.691	(0.049)	-1.811	(0.180)	0.358	(0.062)			86M2	18
N263202	0.771	(0.062)	-0.649	(0.143)	0.405	(0.046)			86M2	19
N264301	0.805	(0.030)	1.004	(0.036)	0.000	(0.000)			86M1	47
N264701	1.463	(0.074)	-0.301	(0.038)	0.208	(0.021)			86M2	39
N266501	0.690	(0.049)	-0.406	(0.120)	0.261	(0.042)			86M3	31
N268801	1.252	(0.075)	1.159	(0.033)	0.091	(0.008)			86M2	48
N268901	1.628	(0.084)	0.430	(0.026)	0.167	(0.012)			86M2	47
N269001	1.319	(0.078)	-0.177	(0.048)	0.271	(0.024)			86M2	22
N270301	0.834	(0.049)	-1.875	(0.121)	0.198	(0.055)			86M1	30
N270302	1.372	(0.059)	-0.045	(0.028)	0.101	(0.014)			86M1	31
N271301	1.216	(0.070)	-0.042	(0.047)	0.249	(0.022)			86M3	32
N278501	0.786	(0.030)	-1.114	(0.040)	0.000	(0.000)			86M1	23
N278502	0.805	(0.029)	-0.839	(0.033)	0.000	(0.000)			86M1	24

Table E-9 (continued)
IRT Parameters for Mathematics Long-term Trend Samples
Age 17

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
N278503	0.641	(0.027)	-1.308	(0.052)	0.000	(0.000)			86M1	25
N278901	0.984	(0.053)	-0.725	(0.070)	0.198	(0.034)			86M2	23
N278902	0.916	(0.053)	-0.502	(0.076)	0.219	(0.033)			86M2	42
N278903	1.253	(0.066)	-0.198	(0.044)	0.221	(0.022)			86M2	44
N278905	0.631	(0.070)	1.193	(0.090)	0.248	(0.026)			86M1	44
N280401	0.526	(0.025)	-1.365	(0.064)	0.000	(0.000)			86M2	30
N281401	0.598	(0.064)	1.415	(0.086)	0.176	(0.023)			86M2	29
N286001	0.662	(0.036)	-1.309	(0.107)	0.164	(0.041)			86M1	19
N286002	1.021	(0.056)	-1.534	(0.086)	0.179	(0.045)			86M1	20
N286301	1.218	(0.062)	-0.845	(0.055)	0.206	(0.031)			86M2	33
N286302	1.211	(0.069)	-0.713	(0.063)	0.282	(0.032)			86M1	22
N286501	1.305	(0.068)	-0.937	(0.055)	0.211	(0.032)			86M2	34
N286502	1.354	(0.063)	-0.448	(0.038)	0.151	(0.021)			86M1	34
N287101	1.125	(0.062)	-0.523	(0.060)	0.232	(0.030)			86M1	29
N287102	1.081	(0.051)	-0.747	(0.053)	0.153	(0.028)			86M2	32
N287301	0.643	(0.024)	0.024	(0.031)	0.000	(0.000)			86M1	45
N287302	0.939	(0.034)	0.972	(0.031)	0.000	(0.000)			86M1	46

Table E-10
IRT Parameters for Mathematics Main Samples
Numbers and Operations, Age 9/Grade 4

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
M010131	0.470	(0.042)	-1.406	(0.216)	0.208	(0.056)			92M8	2
M010231	0.472	(0.050)	-0.213	(0.200)	0.213	(0.052)			92M8	3
M010431	0.719	(0.069)	0.310	(0.101)	0.195	(0.035)			92M8	5
M010531	0.491	(0.092)	2.138	(0.206)	0.164	(0.030)			92M8	6
M010631	0.393	(0.030)	-1.349	(0.117)	0.000	(0.000)			92M8	7
M010831	1.067	(0.077)	0.273	(0.053)	0.140	(0.022)			92M8	9
M011131	1.303	(0.114)	0.395	(0.056)	0.242	(0.024)			92M8	13
M017401	0.506	(0.032)	-2.083	(0.171)	0.185	(0.050)			92M4	1
M017701	1.044	(0.061)	0.545	(0.038)	0.139	(0.015)			92M4	4
M017901	1.361	(0.078)	0.823	(0.028)	0.111	(0.010)			92M4	6
M018201	1.307	(0.104)	1.350	(0.041)	0.154	(0.009)			92M4	9
M018401	1.387	(0.098)	0.967	(0.033)	0.193	(0.011)			92M4	11
M018501	1.251	(0.230)	2.814	(0.242)	0.238	(0.007)			92M4	12
M018601	0.816	(0.136)	2.438	(0.178)	0.154	(0.011)			92M4	13
M020001	1.208	(0.038)	0.122	(0.019)	0.000	(0.000)			92M6	4
M020101	0.919	(0.039)	1.493	(0.046)	0.000	(0.000)			92M6	5
M020501	0.763	(0.032)	1.225	(0.044)	0.000	(0.000)			92M6	9
M021901	0.873	(0.056)	0.096	(0.064)	0.216	(0.025)			92M5	1
M022001	2.483	(0.109)	1.504	(0.030)	0.186	(0.007)			92M5	2
M022301	0.816	(0.051)	-0.124	(0.072)	0.196	(0.028)			92M5	5
M022701	1.064	(0.110)	1.276	(0.056)	0.302	(0.015)			92M5	9
M022901	1.295	(0.119)	1.349	(0.048)	0.261	(0.011)			92M5	12
M023001	1.255	(0.119)	1.468	(0.054)	0.243	(0.011)			92M5	13
M039001	0.760	(0.043)	-1.481	(0.116)	0.199	(0.048)			92M3	1
M039201	0.732	(0.026)	0.032	(0.028)	0.000	(0.000)			92M3	3
M039901	1.045	(0.104)	1.947	(0.077)	0.149	(0.009)			92M3	10
M040201	1.152	(0.060)	1.858	(0.054)	0.000	(0.000)			92M3	13
M040301	0.711	(0.038)	0.457	(0.043)	0.000	(0.000)			92M9	1
M040701	0.772	(0.097)	0.911	(0.096)	0.262	(0.029)			92M9	5
M040901	1.626	(0.075)	0.674	(0.025)	0.000	(0.000)			92M9	7
M041301	0.737	(0.044)	-1.610	(0.080)	0.000	(0.000)			92M12	1
M041401	0.649	(0.068)	0.314	(0.122)	0.210	(0.039)			92M12	2
M041501	1.357	(0.099)	0.358	(0.044)	0.165	(0.019)			92M12	3
M041701	0.607	(0.034)	-0.009	(0.047)	0.000	(0.000)			92M12	5
M041901	0.634	(0.036)	0.515	(0.049)	0.000	(0.000)			92M12	7
M042401	0.944	(0.062)	1.453	(0.065)	0.000	(0.000)			92M12	11
M042601	0.952	(0.081)	0.009	(0.086)	0.256	(0.034)			92M13	1
M042901	0.823	(0.112)	1.240	(0.089)	0.256	(0.025)			92M13	4
M043001	0.787	(0.069)	-0.365	(0.123)	0.260	(0.045)			92M13	5
M043301	0.854	(0.044)	0.896	(0.044)	0.000	(0.000)			92M13	8
M043601	0.273	(0.035)	-4.353	(0.528)	0.000	(0.000)			92M14	1
M044001	0.770	(0.088)	1.034	(0.083)	0.185	(0.025)			92M14	5
M044101	0.627	(0.089)	1.091	(0.120)	0.241	(0.034)			92M14	6
M044301	2.238	(0.162)	2.061	(0.066)	0.055	(0.006)			92M14	9
M044401	0.331	(0.013)	1.376	(0.050)			-1.604	(0.156)	92M14	10
							1.218	(0.197)		
							1.118	(0.182)		
							-0.732	(0.206)		
M044501	0.705	(0.075)	0.655	(0.097)	0.199	(0.032)			92M7	1
M044901	1.357	(0.128)	0.288	(0.062)	0.349	(0.025)			92M7	5
M045001	1.242	(0.143)	1.924	(0.094)	0.145	(0.011)			92M7	6
M045101	0.872	(0.051)	1.333	(0.059)	0.000	(0.000)			92M7	7

BEST COPY AVAILABLE

Table E-10 (continued)
IRT Parameters for Mathematics Main Samples
Numbers and Operations, Age 9/Grade 4

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
M045401	0.316	(0.011)	1.111	(0.042)			-1.039	(0.123)	92M7	10
							-3.244	(0.331)		
							2.878	(0.354)		
							1.405	(0.204)		
M046001	0.663	(0.036)	-0.853	(0.057)	0.000	(0.000)			92M11	1
M046301	1.371	(0.125)	0.827	(0.045)	0.193	(0.017)			92M11	4
M046501	0.881	(0.133)	1.446	(0.094)	0.279	(0.022)			92M11	6
M046801	1.240	(0.059)	0.820	(0.032)	0.000	(0.000)			92M11	9
M046901	1.031	(0.050)	0.693	(0.035)	0.000	(0.000)			92M11	10
M047501	0.590	(0.056)	-0.401	(0.160)	0.211	(0.049)			92M11	16
M048101	0.717	(0.059)	0.130	(0.091)	0.153	(0.032)			92M15	1
M048301	0.681	(0.058)	-0.381	(0.129)	0.209	(0.044)			92M15	3
M048601	0.765	(0.062)	0.268	(0.080)	0.144	(0.029)			92M15	6
M048901	0.873	(0.051)	1.392	(0.060)	0.000	(0.000)			92M15	9
N202831	0.667	(0.063)	-0.158	(0.134)	0.211	(0.045)			92M8	12
N240031	1.343	(0.107)	0.635	(0.042)	0.127	(0.017)			92M8	14
N277903	0.601	(0.026)	-1.148	(0.052)	0.000	(0.000)			92M6	10

Table E-11
IRT Parameters for Mathematics Main Samples
Measurement, Age 9/Grade 4

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
M010731	1.191	(0.130)	0.992	(0.058)	0.260	(0.019)			92M8	8
M010931	0.971	(0.117)	1.245	(0.070)	0.203	(0.020)			92M8	10
M017501	0.582	(0.035)	-1.016	(0.129)	0.166	(0.044)			92M4	2
M018101	1.043	(0.081)	1.124	(0.044)	0.183	(0.013)			92M4	8
M020301	1.020	(0.038)	1.174	(0.033)	0.000	(0.000)			92M6	7
M022601	0.828	(0.113)	1.653	(0.089)	0.318	(0.016)			92M5	8
M022801	1.580	(0.050)	0.178	(0.017)	0.000	(0.000)			92M5	10
M022802	1.290	(0.041)	-0.079	(0.019)	0.000	(0.000)			92M5	11
M023401	1.210	(0.125)	2.041	(0.092)	0.140	(0.008)			92M5	17
M039101	0.556	(0.035)	-0.982	(0.149)	0.184	(0.049)			92M3	2
M039301	0.826	(0.031)	1.043	(0.034)	0.000	(0.000)			92M3	4
M039401	0.872	(0.057)	0.174	(0.065)	0.213	(0.025)			92M3	5
M039501	0.849	(0.045)	-0.120	(0.060)	0.130	(0.025)			92M3	6
M039601	0.737	(0.055)	0.898	(0.055)	0.123	(0.019)			92M3	7
M040461	0.488	(0.015)	-0.724	(0.032)			-0.335	(0.097)	92M9	2
							-1.840	(0.144)		
							2.175	(0.130)		
M040801	1.006	(0.102)	1.054	(0.057)	0.147	(0.019)			92M9	6
M041001	1.410	(0.126)	1.681	(0.065)	0.119	(0.010)			92M9	8
M041601	0.578	(0.096)	1.718	(0.138)	0.177	(0.029)			92M12	4
M042701	0.978	(0.064)	-1.190	(0.091)	0.161	(0.043)			92M13	2
M042801	0.889	(0.061)	-1.348	(0.108)	0.176	(0.048)			92M13	3
M043701	1.214	(0.087)	0.540	(0.041)	0.105	(0.016)			92M14	2
M044601	0.698	(0.037)	0.580	(0.045)	0.000	(0.000)			92M7	2
M047101	1.988	(0.152)	1.797	(0.062)	0.182	(0.009)			92M11	12
M047201	1.291	(0.127)	1.029	(0.048)	0.181	(0.016)			92M11	13
M048201	0.421	(0.040)	-0.751	(0.215)	0.184	(0.053)			92M15	2
M048401	1.008	(0.101)	0.960	(0.059)	0.188	(0.020)			92M15	4
M048501	0.681	(0.066)	0.308	(0.108)	0.179	(0.036)			92M15	5
M048701	1.057	(0.059)	1.376	(0.050)	0.000	(0.000)			92M15	7
M061906	1.308	(0.119)	2.192	(0.105)	0.000	(0.000)			92M10	6

Table E-12
IRT Parameters for Mathematics Main Samples
Geometry, Age 9/Grade 4

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
M011231	0.538	(0.149)	3.012	(0.441)	0.216	(0.025)			92M8	15
M017601	0.239	(0.024)	-0.672	(0.257)	0.244	(0.038)			92M4	3
M018001	0.846	(0.087)	1.516	(0.066)	0.204	(0.015)			92M4	7
M019801	0.356	(0.033)	4.191	(0.361)	0.000	(0.000)			92M6	2
M019901	0.700	(0.025)	-0.422	(0.031)	0.000	(0.000)			92M6	3
M020701	0.498	(0.029)	1.498	(0.081)	0.000	(0.000)			92M6	11
M022201	0.764	(0.028)	0.776	(0.033)	0.000	(0.000)			92M5	4
M022501	0.984	(0.041)	1.488	(0.044)	0.000	(0.000)			92M5	7
M023101	1.213	(0.109)	1.625	(0.056)	0.134	(0.009)			92M5	14
M039801	0.852	(0.162)	2.355	(0.175)	0.350	(0.013)			92M3	9
M041201	0.290	(0.015)	2.352	(0.065)			2.529	(0.114)	92M9	10
							1.256	(0.126)		
							-0.470	(0.194)		
							-3.314	(0.559)		
M041801	1.042	(0.134)	1.687	(0.084)	0.129	(0.014)			92M12	6
M043401	2.328	(0.108)	0.535	(0.018)	0.000	(0.000)			92M13	9
M043402	2.262	(0.108)	0.657	(0.019)	0.000	(0.000)			92M13	10
M043403	1.127	(0.095)	2.221	(0.109)	0.000	(0.000)			92M13	11
M043901	0.649	(0.072)	0.354	(0.130)	0.224	(0.041)			92M14	4
M044801	0.545	(0.058)	-0.142	(0.186)	0.235	(0.053)			92M7	4
M046101	0.634	(0.055)	-0.714	(0.163)	0.226	(0.054)			92M11	2
M046201	1.161	(0.102)	0.764	(0.048)	0.157	(0.018)			92M11	3
M046401	0.883	(0.090)	0.601	(0.078)	0.217	(0.028)			92M11	5
M046701	0.554	(0.085)	[illegible]	(0.134)	0.163	(0.030)			92M11	8
M047401	1.032	(0.079)	-1.487	(0.109)	0.205	(0.052)			92M11	15
M061901	0.646	(0.035)	0.104	(0.044)	0.000	(0.000)			92M10	1
M061902	1.657	(0.071)	0.169	(0.022)	0.000	(0.000)			92M10	2
M061903	1.608	(0.069)	-0.227	(0.024)	0.000	(0.000)			92M10	3
M061904	0.987	(0.063)	1.651	(0.069)	0.000	(0.000)			92M10	4
N214331	0.709	(0.059)	-1.986	(0.170)	0.217	(0.058)			92M8	1

Table E-13
IRT Parameters for Mathematics Main Samples
Data Analysis, Statistics, and Probability, Age 9/Grade 4

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
M017801	0.760	(0.058)	0.857	(0.059)	0.158	(0.020)			92M4	5
M020201	1.152	(0.036)	0.246	(0.020)	0.000	(0.000)			92M6	6
M023301	2.260	(0.114)	1.470	(0.033)	0.183	(0.008)			92M5	16
M040001	1.209	(0.038)	0.487	(0.020)	0.000	(0.000)			92M3	11
M040101	0.477	(0.051)	0.516	(0.165)	0.230	(0.043)			92M3	12
M040601	1.210	(0.082)	0.157	(0.047)	0.127	(0.021)			92M9	4
M041101	2.007	(0.175)	2.033	(0.079)	0.237	(0.010)			92M9	9
M042001	0.900	(0.080)	-0.024	(0.092)	0.218	(0.036)			92M12	8
M042002	0.774	(0.041)	0.220	(0.041)	0.000	(0.000)			92M12	9
M042003	1.247	(0.060)	0.239	(0.029)	0.000	(0.000)			92M12	10
M043101	0.923	(0.094)	0.691	(0.073)	0.231	(0.026)			92M13	6
M043201	0.670	(0.035)	-0.333	(0.045)	0.000	(0.000)			92M13	7
M044701	0.451	(0.079)	1.331	(0.196)	0.245	(0.045)			92M7	3
M045301	1.176	(0.065)	1.130	(0.043)	0.000	(0.000)			92M7	9
M046601	1.274	(0.060)	0.769	(0.031)	0.000	(0.000)			92M11	7
M047001	1.795	(0.134)	1.640	(0.054)	0.152	(0.009)			92M11	11
M047301	1.481	(0.065)	0.072	(0.025)	0.000	(0.000)			92M11	14
M049001	0.488	(0.022)	1.532	(0.045)			0.583	(0.071)	92M15	10
							0.584	(0.093)		
							-0.491	(0.135)		
							-0.676	(0.201)		
M061905	0.893	(0.049)	1.059	(0.049)	0.000	(0.000)			92M10	5
N250231	0.929	(0.085)	0.481	(0.072)	0.189	(0.027)			92M8	11

Table E-14
IRT Parameters for Mathematics Main Samples
Algebra and Functions Age 9/Grade 4

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
M010331	0.630	(0.072)	0.482	(0.136)	0.232	(0.042)			92M8	4
M018301	0.958	(0.065)	1.156	(0.044)	0.101	(0.012)			92M4	10
M018701	0.951	(0.153)	2.992	(0.219)	0.172	(0.008)			92M4	14
M019701	0.751	(0.027)	0.010	(0.027)	0.000	(0.000)			92M6	1
M020401	0.741	(0.028)	0.647	(0.033)	0.000	(0.000)			92M6	8
M022101	0.697	(0.043)	-1.883	(0.144)	0.220	(0.057)			92M5	3
M022401	0.671	(0.066)	0.937	(0.088)	0.260	(0.026)			92M5	6
M023201	1.205	(0.118)	1.475	(0.055)	0.210	(0.011)			92M5	15
M039701	1.117	(0.090)	1.531	(0.047)	0.162	(0.010)			92M3	8
M040501	0.701	(0.065)	-0.079	(0.124)	0.211	(0.044)			92M9	3
M042501	0.967	(0.060)	1.093	(0.056)	0.000	(0.000)			92M12	12
M043501	0.396	(0.018)	1.216	(0.043)			1.161	(0.083)	92M13	12
							-0.569	(0.117)		
							0.367	(0.144)		
							-0.959	(0.179)		
M043801	0.927	(0.043)	0.169	(0.033)	0.000	(0.000)			92M14	3
M044261	0.501	(0.025)	0.802	(0.048)			-0.205	(0.077)	92M14	8
							0.205	(0.091)		
M045201	0.979	(0.152)	2.112	(0.122)	0.186	(0.014)			92M7	8
M048801	1.084	(0.050)	0.622	(0.033)	0.000	(0.000)			92M15	8

Table E-15
IRT Parameters for Mathematics Main Samples
Estimation, Age 9/Grade 4

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
M031101	0.687	(0.045)	-1.545	(0.150)	0.252	(0.057)			92M16	11
M031201	0.274	(0.026)	-0.616	(0.238)	0.251	(0.039)			92M16	12
M031301	0.745	(0.043)	-1.410	(0.116)	0.211	(0.049)			92M16	13
M031401	0.433	(0.065)	1.313	(0.171)	0.268	(0.040)			92M16	14
M031402	0.818	(0.067)	1.209	(0.054)	0.135	(0.016)			92M16	15
M031501	0.572	(0.065)	0.822	(0.123)	0.281	(0.034)			92M16	16
M031601	0.556	(0.038)	-0.641	(0.138)	0.196	(0.044)			92M16	17
M031701	0.552	(0.037)	-1.515	(0.174)	0.234	(0.056)			92M16	18
M031801	1.092	(0.091)	0.785	(0.053)	0.332	(0.017)			92M16	19
M031901	0.126	(0.021)	1.033	(0.432)	0.263	(0.031)			92M16	20
M032001	0.773	(0.056)	0.679	(0.060)	0.156	(0.021)			92M16	1
M032101	0.162	(0.027)	2.908	(0.515)	0.268	(0.027)			92M16	2
M032201	0.417	(0.052)	0.788	(0.190)	0.239	(0.045)			92M16	3
M032301	0.642	(0.066)	1.015	(0.089)	0.232	(0.026)			92M16	4
M032401	0.454	(0.065)	1.720	(0.137)	0.199	(0.031)			92M16	5
M032501	0.590	(0.066)	1.532	(0.090)	0.166	(0.022)			92M16	6
M032601	0.350	(0.050)	1.630	(0.201)	0.254	(0.037)			92M16	7
M032701	0.271	(0.036)	-0.437	(0.342)	0.245	(0.054)			92M16	8
M032901	0.780	(0.122)	2.420	(0.153)	0.164	(0.012)			92M16	10

Table E-16
IRT Parameters for Mathematics Main Samples
Numbers and Operations, Age 13/Grade 8

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
M011131	0.702	(0.043)	-1.151	(0.124)	0.216	(0.048)			92M8	13
M012431	0.892	(0.048)	-0.006	(0.052)	0.131	(0.022)			92M8	3
M012531	0.760	(0.052)	0.812	(0.051)	0.115	(0.018)			92M8	4
M012931	1.142	(0.099)	1.243	(0.045)	0.243	(0.012)			92M8	8
M013431	1.126	(0.068)	0.377	(0.041)	0.166	(0.018)			92M8	15
M013531	0.744	(0.081)	1.762	(0.084)	0.102	(0.014)			92M8	16
M013631	1.556	(0.090)	0.953	(0.026)	0.069	(0.008)			92M8	17
M017401	0.327	(0.031)	-3.969	(0.402)	0.222	(0.060)			92M4	1
M017701	1.035	(0.048)	-0.742	(0.054)	0.145	(0.027)			92M4	4
M017901	1.275	(0.057)	-0.661	(0.042)	0.146	(0.023)			92M4	6
M018201	0.705	(0.035)	-0.580	(0.075)	0.126	(0.029)			92M4	9
M018401	1.113	(0.056)	-0.783	(0.059)	0.199	(0.030)			92M4	11
M018501	1.515	(0.090)	0.600	(0.030)	0.233	(0.012)			92M4	12
M018601	0.658	(0.056)	1.220	(0.065)	0.135	(0.018)			92M4	13
M020001	0.666	(0.024)	-0.209	(0.029)	0.000	(0.000)			92M6	4
M020101	1.196	(0.036)	-0.275	(0.019)	0.000	(0.000)			92M6	5
M020501	0.829	(0.027)	-0.178	(0.025)	0.000	(0.000)			92M6	9
M021901	0.926	(0.046)	-1.098	(0.076)	0.175	(0.036)			92M5	1
M022001	1.182	(0.058)	-0.502	(0.047)	0.190	(0.024)			92M5	2
M022301	0.725	(0.042)	-1.965	(0.137)	0.222	(0.055)			92M5	5
M022701	1.080	(0.057)	-0.505	(0.058)	0.226	(0.027)			92M5	9
M022901	1.311	(0.074)	-0.180	(0.046)	0.291	(0.022)			92M5	12
M023001	1.524	(0.085)	0.042	(0.036)	0.293	(0.017)			92M5	13
M023801	1.316	(0.065)	0.407	(0.028)	0.091	(0.012)			92M5	21
M044501	0.587	(0.045)	-1.418	(0.163)	0.198	(0.051)			92M7	1
M044901	0.794	(0.062)	-1.999	(0.157)	0.225	(0.058)			92M7	5
M045001	2.062	(0.138)	0.102	(0.032)	0.197	(0.018)			92M7	6
M045101	0.734	(0.035)	-0.140	(0.038)	0.000	(0.000)			92M7	7
M046001	0.401	(0.034)	-2.664	(0.214)	0.000	(0.000)			92M11	1
M046301	0.971	(0.070)	-0.816	(0.095)	0.218	(0.042)			92M11	4
M046501	0.979	(0.083)	-0.076	(0.088)	0.293	(0.035)			92M11	6
M046801	0.678	(0.036)	-0.862	(0.056)	0.000	(0.000)			92M11	9
M046901	1.080	(0.053)	-1.067	(0.044)	0.000	(0.000)			92M11	10
M048101	0.527	(0.045)	-2.370	(0.224)	0.211	(0.057)			92M15	1
M048301	0.589	(0.047)	-1.806	(0.183)	0.214	(0.055)			92M15	3
M048601	0.968	(0.071)	-1.550	(0.118)	0.224	(0.052)			92M15	6
M048901	0.739	(0.036)	0.285	(0.038)	0.000	(0.000)			92M15	9
M049101	1.423	(0.112)	1.331	(0.041)	0.060	(0.008)			92M15	10
M049901	0.848	(0.060)	-0.284	(0.082)	0.148	(0.034)			92M3	1
M050001	1.340	(0.075)	-0.188	(0.040)	0.095	(0.020)			92M3	2
M050101	1.109	(0.077)	0.809	(0.040)	0.065	(0.014)			92M3	3
M050301	1.374	(0.123)	1.338	(0.049)	0.177	(0.012)			92M3	5
M051101	0.347	(0.017)	1.723	(0.049)			1.495	(0.092)	92M3	13
							0.831	(0.106)		
							-0.165	(0.140)		
							-2.161	(0.285)		
M051201	0.537	(0.034)	-1.759	(0.106)	0.000	(0.000)			92M13	1
M051501	0.833	(0.125)	1.859	(0.113)	0.161	(0.017)			92M13	4
M051601	0.864	(0.041)	-0.689	(0.041)	0.000	(0.000)			92M13	5
M051901	2.181	(0.132)	0.864	(0.024)	0.077	(0.009)			92M13	8
M052401	0.894	(0.044)	0.933	(0.041)	0.000	(0.000)			92M9	2
M052901	0.734	(0.036)	0.189	(0.038)	0.000	(0.000)			92M9	7
M053001	0.910	(0.050)	1.278	(0.052)	0.000	(0.000)			92M9	8
M053601	0.835	(0.065)	0.101	(0.081)	0.168	(0.032)			92M12	2

Table E-16 (continued)
IRT Parameters for Mathematics Main Samples
Numbers and Operations, Age 13/Grade 8

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
M053701	0.821	(0.102)	0.779	(0.100)	0.339	(0.031)			92M12	3
M053901	0.878	(0.073)	0.881	(0.056)	0.102	(0.019)			92M12	5
M054701	0.507	(0.046)	-0.987	(0.203)	0.191	(0.058)			92M14	1
M054801	0.681	(0.034)	0.236	(0.040)	0.000	(0.000)			92M14	2
M055201	1.082	(0.046)	0.141	(0.028)	0.000	(0.000)			92M14	6
M055501	0.485	(0.025)	1.744	(0.052)			1.022	(0.073)	92M14	9
							0.518	(0.094)		
							-1.628	(0.209)		
							0.088	(0.295)		
N202831	0.642	(0.042)	-1.957	(0.175)	0.252	(0.063)			92M8	12

Table E-17
IRT Parameters for Mathematics Main Samples
Measurement, Age 13/Grade 8

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
M012331	0.826	(0.041)	-1.262	(0.086)	0.155	(0.041)			92M8	2
M013331	1.057	(0.056)	-0.988	(0.070)	0.172	(0.038)			92M8	14
M017501	0.556	(0.035)	-1.915	(0.168)	0.218	(0.056)			92M4	2
M018101	0.791	(0.046)	-0.197	(0.074)	0.177	(0.030)			92M4	8
M019101	1.607	(0.097)	1.896	(0.050)	0.174	(0.006)			92M4	18
M019201	1.555	(0.092)	1.820	(0.047)	0.150	(0.006)			92M4	19
M020301	0.985	(0.030)	-0.276	(0.022)	0.000	(0.000)			92M6	7
M022601	0.665	(0.062)	0.600	(0.099)	0.279	(0.030)			92M5	8
M022801	1.811	(0.055)	-0.528	(0.015)	0.000	(0.000)			92M5	10
M022802	1.690	(0.054)	-0.819	(0.018)	0.000	(0.000)			92M5	11
M023401	1.094	(0.077)	0.265	(0.056)	0.315	(0.022)			92M5	17
M023701	0.697	(0.028)	0.948	(0.040)	0.000	(0.000)			92M5	20
M044601	0.701	(0.035)	-0.497	(0.043)	0.000	(0.000)			92M7	2
M047101	1.119	(0.085)	0.500	(0.051)	0.164	(0.021)			92M11	12
M047201	0.957	(0.069)	-0.385	(0.083)	0.200	(0.036)			92M11	13
M047901	0.916	(0.056)	1.833	(0.074)	0.000	(0.000)			92M11	18
M048201	0.503	(0.044)	-0.829	(0.187)	0.214	(0.053)			92M15	2
M048401	0.985	(0.065)	-0.710	(0.079)	0.173	(0.037)			92M15	4
M048501	0.782	(0.059)	-0.963	(0.128)	0.227	(0.051)			92M15	5
M048701	0.649	(0.033)	-0.288	(0.043)	0.000	(0.000)			92M15	7
M049201	0.542	(0.069)	0.822	(0.146)	0.219	(0.041)			92M15	11
M049501	0.899	(0.102)	1.478	(0.074)	0.155	(0.017)			92M15	14
M050501	0.646	(0.068)	0.358	(0.126)	0.219	(0.040)			92M3	7
M050901	0.795	(0.046)	1.616	(0.070)	0.000	(0.000)			92M3	11
M051301	0.451	(0.037)	-2.555	(0.192)	0.000	(0.000)			92M13	2
M052201	0.453	(0.020)	1.494	(0.045)			0.937	(0.067)	92M13	11
							-0.672	(0.110)		
							-1.082	(0.202)		
							0.817	(0.224)		
M052301	1.585	(0.130)	1.119	(0.040)	0.182	(0.012)			92M9	1
M054001	0.750	(0.055)	2.456	(0.130)	0.000	(0.000)			92M12	6
M055101	0.628	(0.114)	2.197	(0.171)	0.204	(0.022)			92M14	5
M055401	0.830	(0.088)	1.421	(0.075)	0.118	(0.017)			92M14	8
M061907	0.981	(0.049)	1.219	(0.046)	0.000	(0.000)			92M10	5
M061908	0.839	(0.074)	2.764	(0.165)	0.000	(0.000)			92M10	6

Table E-18
IRT Parameters for Mathematics Main Samples
Geometry, Age 13/Grade 8

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
M012731	0.628	(0.064)	1.456	(0.078)	0.162	(0.021)			92M8	6
M012831	1.113	(0.065)	0.759	(0.034)	0.125	(0.013)			92M8	7
M017601	0.605	(0.042)	-1.018	(0.166)	0.282	(0.054)			92M4	3
M018001	0.795	(0.057)	0.313	(0.073)	0.243	(0.026)			92M4	7
M019001	0.816	(0.059)	0.937	(0.051)	0.148	(0.017)			92M4	17
M019601	0.615	(0.066)	1.755	(0.088)	0.122	(0.018)			92M4	21
M019801	0.890	(0.029)	-0.514	(0.025)	0.000	(0.000)			92M6	2
M019901	0.686	(0.027)	-1.292	(0.047)	0.000	(0.000)			92M6	3
M020901	0.576	(0.034)	0.994	(0.062)	0.000	(0.000)			92M6	11
M021001	0.892	(0.029)	0.396	(0.024)	0.000	(0.000)			92M6	12
M021301	1.463	(0.044)	0.178	(0.017)	0.000	(0.000)			92M6	15
M021302	1.325	(0.041)	0.001	(0.018)	0.000	(0.000)			92M6	16
M022201	0.690	(0.035)	-0.655	(0.047)	0.000	(0.000)			92M5	4
M022501	0.928	(0.029)	-0.137	(0.022)	0.000	(0.000)			92M5	7
M023101	1.375	(0.069)	0.237	(0.031)	0.162	(0.014)			92M5	14
M044801	0.675	(0.055)	-1.567	(0.175)	0.264	(0.061)			92M7	4
M045601	0.315	(0.026)	0.366	(0.083)	0.000	(0.000)			92M7	10
M045901	0.590	(0.028)	1.868	(0.044)			0.941	(0.058)	92M7	13
							0.528	(0.081)		
							-0.544	(0.137)		
							-0.925	(0.243)		
M046101	0.854	(0.069)	-1.937	(0.151)	0.254	(0.060)			92M11	2
M046201	1.232	(0.099)	-0.178	(0.070)	0.303	(0.031)			92M11	3
M046401	0.795	(0.069)	-0.523	(0.129)	0.277	(0.048)			92M11	5
M046701	0.995	(0.092)	0.317	(0.077)	0.288	(0.029)			92M11	8
M048001	1.272	(0.128)	1.607	(0.062)	0.086	(0.010)			92M11	19
M049301	1.065	(0.117)	1.591	(0.068)	0.120	(0.013)			92M15	12
M049701	1.479	(0.119)	0.944	(0.038)	0.127	(0.013)			92M15	16
M051001	0.655	(0.038)	1.185	(0.064)	0.000	(0.000)			92M3	12
M051801	1.153	(0.141)	1.946	(0.100)	0.253	(0.012)			92M13	7
M052001	0.834	(0.102)	1.557	(0.083)	0.155	(0.018)			92M13	9
M052601	0.832	(0.075)	0.525	(0.076)	0.182	(0.028)			92M9	4
M054101	0.886	(0.040)	0.118	(0.033)	0.000	(0.000)			92M12	7
M054201	1.053	(0.148)	1.738	(0.088)	0.226	(0.015)			92M12	8
M055301	0.470	(0.158)	4.695	(1.085)	0.207	(0.015)			92M14	7
M061901	0.660	(0.036)	-0.952	(0.057)	0.000	(0.000)			92M10	1
M061902	1.810	(0.083)	-0.758	(0.025)	0.000	(0.000)			92M10	4
M061903	1.470	(0.071)	-1.118	(0.034)	0.000	(0.000)			92M10	2
M061904	1.059	(0.046)	0.330	(0.029)	0.000	(0.000)			92M10	3

BEST COPY AVAILABLE

Table E-19
IRT Parameters for Mathematics Main Samples
Data Analysis, Statistics, and Probability, Age 13/Grade 8

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
M012631	1.455	(0.092)	0.817	(0.030)	0.211	(0.011)			92M8	5
M013031	1.214	(0.053)	1.554	(0.039)	0.000	(0.000)			92M8	9
M013131	0.958	(0.041)	1.502	(0.044)	0.000	(0.000)			92M8	10
M017801	1.120	(0.061)	-0.274	(0.054)	0.256	(0.024)			92M4	5
M018901	1.578	(0.119)	2.002	(0.069)	0.180	(0.007)			92M4	16
M020201	0.641	(0.027)	-1.786	(0.064)	0.000	(0.000)			92M6	6
M020801	1.192	(0.054)	1.678	(0.043)	0.000	(0.000)			92M6	10
M021101	0.986	(0.031)	0.301	(0.022)	0.000	(0.000)			92M6	13
M023301	1.881	(0.092)	-0.366	(0.029)	0.205	(0.017)			92M5	16
M023501	1.379	(0.083)	0.951	(0.028)	0.120	(0.010)			92M5	18
M023601	1.015	(0.052)	-0.107	(0.049)	0.138	(0.021)			92M5	19
M044701	1.013	(0.120)	0.693	(0.086)	0.407	(0.025)			92M7	3
M045301	0.984	(0.041)	-0.141	(0.031)	0.000	(0.000)			92M7	9
M045861	0.444	(0.015)	-0.591	(0.027)			1.369	(0.109)	92M7	12
							-0.613	(0.092)		
							-0.574	(0.097)		
							-0.181	(0.083)		
M046601	1.144	(0.051)	-0.785	(0.034)	0.000	(0.000)			92M11	7
M047001	0.987	(0.078)	0.413	(0.063)	0.175	(0.025)			92M11	11
M047301	1.022	(0.057)	-1.538	(0.059)	0.000	(0.000)			92M11	14
M047801	1.708	(0.130)	0.900	(0.032)	0.122	(0.012)			92M11	17
M049801	0.772	(0.069)	2.465	(0.156)	0.000	(0.000)			92M15	17
M050261	0.859	(0.029)	-1.471	(0.026)			0.979	(0.113)	92M3	4
							-0.772	(0.090)		
							-0.071	(0.085)		
							-0.136	(0.049)		
M050401	0.756	(0.071)	0.199	(0.108)	0.209	(0.039)			92M3	6
M051401	0.911	(0.057)	-0.699	(0.080)	0.139	(0.037)			92M13	3
M052701	1.306	(0.113)	0.764	(0.047)	0.214	(0.018)			92M9	5
M052801	0.625	(0.226)	4.360	(1.131)	0.185	(0.012)			92M9	6
M053101	0.669	(0.024)	1.384	(0.028)			0.205	(0.052)	92M9	9
							-1.249	(0.135)		
							1.221	(0.148)		
							-0.178	(0.113)		
M053801	0.871	(0.167)	2.114	(0.158)	0.225	(0.016)			92M12	4
M054901	1.094	(0.130)	1.965	(0.102)	0.156	(0.011)			92M14	3
M061905	0.533	(0.031)	-0.184	(0.053)	0.000	(0.000)			92M10	7

Table E-20
IRT Parameters for Mathematics Main Samples
Algebra and Functions, Age 13/Grade 8

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
M012231	0.497	(0.036)	-3.078	(0.214)	0.139	(0.049)			92M8	1
M013231	1.280	(0.091)	1.802	(0.054)	0.129	(0.007)			92M8	11
M013731	1.204	(0.111)	1.441	(0.051)	0.146	(0.011)			92M8	18
M018301	1.018	(0.055)	-0.153	(0.054)	0.204	(0.024)			92M4	10
M018701	1.495	(0.090)	0.482	(0.032)	0.263	(0.013)			92M4	14
M018801	0.850	(0.077)	1.083	(0.059)	0.260	(0.018)			92M4	15
M019301	1.226	(0.102)	1.433	(0.044)	0.186	(0.010)			92M4	20
M019701	0.492	(0.023)	-1.526	(0.069)	0.000	(0.000)			92M6	1
M020401	0.663	(0.024)	0.046	(0.029)	0.000	(0.000)			92M6	8
M021201	1.061	(0.034)	0.679	(0.024)	0.000	(0.000)			92M6	14
M022101	0.808	(0.053)	-2.437	(0.148)	0.262	(0.061)			92M5	3
M022401	1.141	(0.071)	-0.515	(0.071)	0.357	(0.030)			92M5	6
M023201	1.346	(0.070)	-0.052	(0.038)	0.212	(0.019)			92M5	15
M045201	0.877	(0.079)	0.679	(0.068)	0.179	(0.024)			92M7	8
M045701	1.158	(0.048)	0.273	(0.027)	0.000	(0.000)			92M7	11
M047601	1.494	(0.095)	0.558	(0.032)	0.079	(0.013)			92M11	15
M047701	1.210	(0.134)	1.726	(0.080)	0.248	(0.013)			92M11	16
M048801	1.026	(0.046)	-0.776	(0.036)	0.000	(0.000)			92M15	8
M049401	1.227	(0.119)	1.487	(0.054)	0.105	(0.011)			92M15	13
M049601	0.942	(0.070)	0.007	(0.073)	0.156	(0.031)			92M15	15
M050601	1.288	(0.094)	0.824	(0.038)	0.094	(0.013)			92M3	8
M050701	1.547	(0.095)	0.187	(0.035)	0.112	(0.017)			92M3	9
M050801	0.744	(0.036)	0.258	(0.038)	0.000	(0.000)			92M3	10
M051701	1.027	(0.083)	0.657	(0.053)	0.142	(0.020)			92M13	6
M052101	1.038	(0.046)	0.535	(0.031)	0.000	(0.000)			92M13	10
M052501	0.827	(0.094)	1.490	(0.075)	0.126	(0.018)			92M9	3
M053501	1.808	(0.108)	-0.263	(0.035)	0.123	(0.020)			92M12	1
M054301	0.404	(0.015)	1.846	(0.056)			-1.181	(0.109)	92M12	9
							0.391	(0.165)		
							-2.119	(0.377)		
							2.909	(0.386)		
M055001	0.705	(0.098)	1.987	(0.125)	0.095	(0.017)			92M14	4

Table E-21
IRT Parameters for Mathematics Main Samples Estimation, Age 13/Grade 8

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
M032001	0.690	(0.055)	0.203	(0.095)	0.282	(0.030)			92M16	1
M032201	0.601	(0.043)	-1.289	(0.188)	0.340	(0.055)			92M16	3
M032301	0.696	(0.039)	-0.649	(0.093)	0.156	(0.035)			92M16	4
M032401	0.481	(0.058)	0.381	(0.187)	0.339	(0.043)			92M16	5
M032501	0.355	(0.033)	0.133	(0.183)	0.291	(0.036)			92M16	6
M032601	1.040	(0.079)	0.229	(0.066)	0.413	(0.022)			92M16	7
M032701	0.550	(0.039)	-1.602	(0.210)	0.346	(0.058)			92M16	8
M032801	2.218	(0.104)	0.953	(0.022)	0.311	(0.009)			92M16	9
M032901	1.116	(0.068)	0.282	(0.046)	0.268	(0.019)			92M16	10
M033001	0.921	(0.078)	1.147	(0.046)	0.179	(0.015)			92M16	11
M033101	0.579	(0.049)	0.277	(0.111)	0.245	(0.033)			92M16	12
M033201	0.574	(0.055)	-0.195	(0.173)	0.371	(0.043)			92M16	13
M033301	1.212	(0.062)	0.624	(0.027)	0.091	(0.011)			92M16	14
M033401	0.474	(0.067)	1.046	(0.149)	0.315	(0.036)			92M16	15
M033501	0.772	(0.082)	0.674	(0.088)	0.424	(0.024)			92M16	16
M033601	0.556	(0.071)	1.904	(0.111)	0.120	(0.019)			92M16	17
M033701	1.191	(0.098)	1.192	(0.038)	0.198	(0.011)			92M16	18
M033801	1.293	(0.076)	0.395	(0.037)	0.257	(0.016)			92M16	19
M033901	0.773	(0.084)	1.114	(0.066)	0.292	(0.020)			92M16	20
M034001	0.485	(0.095)	2.468	(0.224)	0.211	(0.023)			92M16	21
M034101	1.394	(0.086)	1.290	(0.032)	0.114	(0.008)			92M16	22

Table E-22
IRT Parameters for Mathematics Main Samples
Numbers and Operations, Age 17/Grade 12

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
M011531	2.839	(0.100)	1.272	(0.020)	0.095	(0.005)			92M8	15
M012431	0.554	(0.034)	-1.748	(0.161)	0.187	(0.052)			92M8	3
M012531	0.654	(0.039)	-0.043	(0.075)	0.118	(0.027)			92M8	4
M012931	1.127	(0.073)	0.526	(0.042)	0.228	(0.017)			92M8	8
M017401	0.237	(0.032)	-6.025	(0.825)	0.231	(0.062)			92M4	1
M017701	0.878	(0.047)	-1.610	(0.096)	0.188	(0.044)			92M4	4
M017901	1.262	(0.064)	-1.398	(0.062)	0.182	(0.035)			92M4	6
M018201	0.586	(0.035)	-1.525	(0.142)	0.192	(0.048)			92M4	9
M018401	1.079	(0.060)	-1.335	(0.084)	0.243	(0.042)			92M4	11
M018501	1.847	(0.099)	-0.012	(0.028)	0.242	(0.016)			92M4	12
M018601	0.527	(0.045)	0.658	(0.099)	0.151	(0.030)			92M4	13
M020501	0.808	(0.030)	-1.025	(0.037)	0.000	(0.000)			92M6	9
M021401	0.804	(0.028)	-0.286	(0.027)	0.000	(0.000)			92M6	1
M021701	1.443	(0.049)	0.578	(0.019)	0.000	(0.000)			92M6	15
M021901	0.823	(0.045)	-1.997	(0.108)	0.177	(0.045)			92M5	1
M022001	0.852	(0.044)	-1.620	(0.090)	0.162	(0.040)			92M5	2
M023901	0.403	(0.021)	-0.919	(0.063)	0.000	(0.000)			92M5	3
M025101	1.421	(0.113)	1.878	(0.066)	0.244	(0.008)			92M5	15
M025302	0.794	(0.035)	1.257	(0.047)	0.000	(0.000)			92M5	17
M052401	0.660	(0.037)	0.255	(0.043)	0.000	(0.000)			92M9	2
M053601	0.743	(0.055)	-0.709	(0.108)	0.161	(0.040)			92M12	2
M053701	1.374	(0.132)	0.219	(0.064)	0.395	(0.026)			92M12	3
M053901	1.304	(0.091)	-0.166	(0.052)	0.175	(0.026)			92M12	5
M054701	0.613	(0.051)	-1.835	(0.185)	0.190	(0.055)			92M14	1
M054801	0.381	(0.029)	-0.601	(0.083)	0.000	(0.000)			92M14	2
M055201	0.934	(0.045)	-0.586	(0.040)	0.000	(0.000)			92M14	6
M056501	1.537	(0.141)	1.372	(0.054)	0.230	(0.013)			92M3	8
M056601	1.064	(0.074)	0.632	(0.043)	0.074	(0.016)			92M3	9
M056801	2.106	(0.106)	1.020	(0.025)	0.000	(0.000)			92M3	10
M057101	1.527	(0.139)	2.228	(0.098)	0.000	(0.000)			92M3	14
M057201	0.460	(0.038)	-2.790	(0.211)	0.000	(0.000)			92M7	1
M057901	1.083	(0.116)	1.315	(0.061)	0.155	(0.017)			92M7	8
M058401	0.588	(0.079)	3.270	(0.363)	0.000	(0.000)			92M7	13
M058601	1.211	(0.087)	0.010	(0.052)	0.168	(0.025)			92M11	2
M059001	1.160	(0.080)	0.480	(0.042)	0.101	(0.017)			92M11	5
M059201	0.786	(0.054)	-0.952	(0.102)	0.152	(0.040)			92M11	7
M059601	1.249	(0.133)	1.500	(0.059)	0.113	(0.012)			92M11	11
M060101	0.695	(0.080)	0.834	(0.098)	0.203	(0.031)			92M13	3
M060801	0.537	(0.034)	-1.234	(0.084)	0.000	(0.000)			92M15	1
M061001	0.688	(0.037)	0.237	(0.041)	0.000	(0.000)			92M15	3
M061201	1.732	(0.146)	1.836	(0.069)	0.204	(0.010)			92M15	5
M062101	0.590	(0.037)	0.825	(0.060)	0.000	(0.000)			92M10	7
N202831	0.717	(0.046)	-1.935	(0.155)	0.249	(0.059)			92M8	12

Table E-23
IRT Parameters for Mathematics Main Samples
Measurement, Age 17/Grade 12

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
M011331	2.131	(0.134)	0.640	(0.023)	0.244	(0.011)			92M8	13
M011931	2.695	(0.110)	1.835	(0.033)	0.000	(0.000)			92M8	19
M012331	0.781	(0.043)	-1.926	(0.120)	0.191	(0.053)			92M8	2
M017501	0.430	(0.032)	-2.871	(0.254)	0.222	(0.059)			92M4	2
M018101	0.552	(0.036)	-0.825	(0.140)	0.187	(0.044)			92M4	8
M019101	0.972	(0.078)	1.118	(0.046)	0.170	(0.014)			92M4	18
M019201	1.551	(0.099)	1.152	(0.030)	0.149	(0.009)			92M4	19
M020301	1.047	(0.036)	-1.113	(0.031)	0.000	(0.000)			92M6	7
M024001	0.652	(0.039)	-2.053	(0.151)	0.202	(0.057)			92M5	4
M024301	1.148	(0.057)	0.402	(0.030)	0.089	(0.013)			92M5	7
M024601	0.721	(0.051)	0.238	(0.075)	0.174	(0.027)			92M5	10
M024701	2.688	(0.098)	0.793	(0.013)	0.000	(0.000)			92M5	11
M024801	2.036	(0.118)	0.191	(0.027)	0.304	(0.014)			92M5	12
M052301	1.746	(0.129)	0.464	(0.034)	0.165	(0.016)			92M9	1
M054001	1.632	(0.083)	1.018	(0.028)	0.000	(0.000)			92M12	6
M055101	1.194	(0.115)	0.993	(0.050)	0.180	(0.017)			92M14	5
M055401	1.035	(0.095)	0.578	(0.060)	0.201	(0.023)			92M14	8
M055801	0.586	(0.046)	-1.786	(0.176)	0.187	(0.054)			92M3	1
M057301	1.012	(0.076)	-0.435	(0.086)	0.209	(0.037)			92M7	2
M057601	1.312	(0.085)	0.308	(0.038)	0.100	(0.017)			92M7	5
M058201	2.057	(0.113)	1.653	(0.041)	0.000	(0.000)			92M7	11
M058501	0.805	(0.066)	0.499	(0.066)	0.117	(0.024)			92M11	1
M059801	1.944	(0.132)	1.917	(0.059)	0.000	(0.000)			92M11	14
M060201	0.693	(0.041)	1.145	(0.061)	0.000	(0.000)			92M13	4
M060501	1.061	(0.056)	1.002	(0.040)	0.000	(0.000)			92M13	7
M061601	1.222	(0.259)	2.563	(0.255)	0.273	(0.011)			92M15	9
M061907	0.812	(0.044)	0.960	(0.047)	0.000	(0.000)			92M10	3
M061908	0.931	(0.061)	1.750	(0.076)	0.000	(0.000)			92M10	4

Table E-24
IRT Parameters for Mathematics Main Samples
Geometry, Age 17/Grade 12

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
M011731	1.491	(0.091)	0.543	(0.033)	0.196	(0.014)			92M8	17
M012731	0.994	(0.054)	0.061	(0.049)	0.166	(0.021)			92M8	6
M012831	1.704	(0.078)	-0.481	(0.030)	0.139	(0.017)			92M8	7
M017601	0.513	(0.033)	-1.998	(0.183)	0.202	(0.056)			92M4	3
M018001	0.721	(0.049)	-0.432	(0.110)	0.229	(0.040)			92M4	7
M019001	0.927	(0.049)	-0.142	(0.055)	0.145	(0.024)			92M4	17
M020901	0.767	(0.039)	0.456	(0.040)	0.000	(0.000)			92M6	11
M021001	0.916	(0.031)	-0.715	(0.028)	0.000	(0.000)			92M6	12
M021801	1.556	(0.074)	1.489	(0.032)	0.000	(0.000)			92M6	17
M024101	1.188	(0.060)	-0.261	(0.044)	0.171	(0.022)			92M5	5
M024401	1.553	(0.103)	0.922	(0.031)	0.219	(0.011)			92M5	8
M052601	0.677	(0.054)	-0.026	(0.099)	0.142	(0.035)			92M9	4
M053301	1.377	(0.077)	1.326	(0.040)	0.000	(0.000)			92M9	8
M055301	1.550	(0.154)	1.014	(0.047)	0.279	(0.016)			92M14	7
M055601	1.446	(0.074)	1.071	(0.033)	0.000	(0.000)			92M14	9
M056001	1.119	(0.076)	-0.334	(0.064)	0.163	(0.031)			92M3	3
M056701	1.795	(0.120)	1.505	(0.044)	0.126	(0.009)			92M3	11
M057461	0.340	(0.015)	-1.270	(0.044)			2.176	(0.217)	92M7	3
							-0.710	(0.131)		
							-0.510	(0.117)		
							-0.956	(0.096)		
M057801	2.252	(0.133)	1.287	(0.035)	0.215	(0.011)			92M7	7
M058101	1.524	(0.136)	1.056	(0.040)	0.142	(0.013)			92M7	10
M058701	1.290	(0.098)	-0.221	(0.064)	0.249	(0.031)			92M11	3
M058901	0.914	(0.043)	0.411	(0.034)	0.000	(0.000)			92M11	4
M059101	1.411	(0.127)	0.446	(0.054)	0.311	(0.022)			92M11	6
M060001	1.368	(0.070)	1.205	(0.036)	0.000	(0.000)			92M13	2
M060601	1.051	(0.059)	1.330	(0.049)	0.000	(0.000)			92M13	8
M060901	0.925	(0.077)	-0.122	(0.089)	0.227	(0.037)			92M15	2
M061104	0.658	(0.063)	-0.483	(0.165)	0.332	(0.050)			92M15	4
M061161	0.440	(0.020)	-1.772	(0.045)			2.189	(0.296)	92M15	4
							0.059	(0.125)		
							-1.018	(0.090)		
							-1.231	(0.072)		
M061301	0.693	(0.076)	0.679	(0.101)	0.193	(0.034)			92M15	6
M061901	0.683	(0.041)	-1.441	(0.079)	0.000	(0.000)			92M10	1
M061904	0.635	(0.035)	-0.074	(0.045)	0.000	(0.000)			92M10	2
M062401	0.230	(0.013)	2.864	(0.099)			2.032	(0.152)	92M10	10
							-0.950	(0.256)		
							3.455	(0.293)		
							-4.537	(0.685)		

Table E-25
IRT Parameters for Mathematics Main Samples
Data Analysis, Statistics, and Probability, Age 17/Grade 12

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
M011631	0.857	(0.057)	0.073	(0.063)	0.189	(0.025)			92M8	16
M012631	1.334	(0.082)	-0.263	(0.051)	0.337	(0.022)			92M8	5
M013031	0.943	(0.036)	0.853	(0.029)	0.000	(0.000)			92M8	9
M013131	0.690	(0.030)	1.009	(0.042)	0.000	(0.000)			92M8	10
M017801	1.257	(0.073)	-0.772	(0.066)	0.322	(0.030)			92M4	5
M018901	1.428	(0.136)	2.117	(0.088)	0.202	(0.007)			92M4	16
M020201	0.672	(0.031)	-2.230	(0.082)	0.000	(0.000)			92M6	6
M020801	0.979	(0.040)	1.154	(0.035)	0.000	(0.000)			92M6	10
M021101	0.900	(0.031)	-0.566	(0.029)	0.000	(0.000)			92M6	13
M021501	1.228	(0.038)	-0.378	(0.021)	0.000	(0.000)			92M6	2
M021502	0.961	(0.033)	-1.093	(0.034)	0.000	(0.000)			92M6	3
M023501	1.464	(0.087)	0.392	(0.029)	0.147	(0.014)			92M5	18
M023601	0.809	(0.054)	-0.658	(0.097)	0.204	(0.037)			92M5	19
M024501	0.357	(0.039)	0.575	(0.186)	0.273	(0.037)			92M5	9
M052701	1.166	(0.085)	-0.230	(0.064)	0.186	(0.029)			92M9	5
M052801	0.300	(0.092)	4.975	(1.141)	0.202	(0.027)			92M9	6
M053401	0.276	(0.015)	0.954	(0.049)			0.833	(0.135)	92M9	9
							0.297	(0.148)		
							0.274	(0.161)		
							-1.405	(0.208)		
M053801	0.413	(0.127)	3.550	(0.684)	0.276	(0.027)			92M12	4
M054901	0.763	(0.148)	2.021	(0.176)	0.213	(0.020)			92M14	3
M055901	0.655	(0.054)	-0.555	(0.132)	0.184	(0.045)			92M3	2
M057001	1.512	(0.115)	1.798	(0.067)	0.000	(0.000)			92M3	13
M058301	1.570	(0.120)	1.742	(0.064)	0.000	(0.000)			92M7	12
M059701	0.400	(0.121)	3.302	(0.611)	0.289	(0.031)			92M11	12
M059702	0.837	(0.057)	1.487	(0.074)	0.000	(0.000)			92M11	13
M059901	0.878	(0.071)	-0.433	(0.103)	0.212	(0.040)			92M13	1
M060301	0.575	(0.058)	-0.435	(0.181)	0.231	(0.054)			92M13	5
M060401	0.852	(0.041)	-0.197	(0.037)	0.000	(0.000)			92M13	6
M061401	0.937	(0.136)	1.839	(0.113)	0.125	(0.015)			92M15	7
M061905	0.464	(0.031)	-1.008	(0.085)	0.000	(0.000)			92M10	5

Table E-26
IRT Parameters for Mathematics Main Samples
Algebra and Functions, Age 17/Grade 12

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
M011431	1.370	(0.097)	0.893	(0.035)	0.244	(0.013)			92M8	14
M011831	2.336	(0.110)	1.593	(0.034)	0.218	(0.008)			92M8	18
M012031	1.304	(0.055)	0.976	(0.029)	0.000	(0.000)			92M8	20
M012131	1.862	(0.120)	1.436	(0.042)	0.166	(0.009)			92M8	21
M012231	0.514	(0.043)	-3.565	(0.278)	0.200	(0.057)			92M8	1
M013231	1.215	(0.082)	0.976	(0.035)	0.167	(0.012)			92M8	11
M018301	0.759	(0.039)	-1.085	(0.091)	0.149	(0.039)			92M4	10
M018701	1.371	(0.066)	-0.642	(0.043)	0.160	(0.024)			92M4	14
M018801	0.974	(0.054)	-0.324	(0.064)	0.207	(0.028)			92M4	15
M019301	1.843	(0.097)	0.266	(0.026)	0.191	(0.013)			92M4	20
M019401	2.290	(0.121)	1.714	(0.041)	0.292	(0.008)			92M4	21
M019501	1.515	(0.084)	1.554	(0.037)	0.083	(0.006)			92M4	22
M020401	0.697	(0.025)	-0.388	(0.031)	0.000	(0.000)			92M6	8
M021201	0.778	(0.027)	-0.374	(0.030)	0.000	(0.000)			92M6	14
M021601	1.351	(0.065)	1.756	(0.043)	0.000	(0.000)			92M6	4
M021602	1.964	(0.073)	1.394	(0.023)	0.000	(0.000)			92M6	5
M021702	0.743	(0.028)	0.083	(0.030)	0.000	(0.000)			92M6	16
M024201	1.090	(0.052)	0.390	(0.032)	0.077	(0.013)			92M5	6
M025001	1.767	(0.083)	-0.473	(0.030)	0.144	(0.018)			92M5	14
M025201	1.987	(0.111)	1.760	(0.044)	0.191	(0.007)			92M5	16
M025401	1.766	(0.245)	0.782	(0.064)	0.411	(0.021)			92M5	20
M052501	1.331	(0.093)	0.140	(0.047)	0.164	(0.021)			92M9	3
M053201	1.304	(0.125)	0.814	(0.051)	0.257	(0.019)			92M9	7
M053501	1.118	(0.074)	-1.252	(0.084)	0.173	(0.043)			92M12	1
M054401	1.364	(0.068)	1.036	(0.033)	0.000	(0.000)			92M12	7
M054501	1.853	(0.109)	1.818	(0.049)	0.000	(0.000)			92M12	8
M054601	0.566	(0.029)	2.138	(0.069)			-0.461	(0.091)	92M12	9
							-0.072	(0.178)		
							0.550	(0.236)		
							-0.017	(0.294)		
M055001	0.506	(0.088)	2.103	(0.186)	0.135	(0.027)			92M14	4
M055701	0.556	(0.031)	2.118	(0.075)			0.248	(0.071)	92M14	10
							-1.410	(0.221)		
							0.058	(0.366)		
							1.105	(0.375)		
M056101	1.014	(0.103)	0.722	(0.066)	0.243	(0.024)			92M3	4
M056201	1.205	(0.157)	1.246	(0.066)	0.306	(0.018)			92M3	5
M056301	1.580	(0.097)	0.257	(0.032)	0.092	(0.014)			92M3	6
M056401	1.567	(0.160)	0.839	(0.048)	0.319	(0.017)			92M3	7
M056901	2.306	(0.108)	1.436	(0.030)	0.000	(0.000)			92M3	12
M057501	1.043	(0.045)	-0.369	(0.032)	0.000	(0.000)			92M7	4
M057701	1.432	(0.101)	0.866	(0.035)	0.087	(0.012)			92M7	6
M058001	1.507	(0.067)	0.588	(0.025)	0.000	(0.000)			92M7	9
M059301	1.092	(0.084)	0.650	(0.047)	0.116	(0.018)			92M11	8
M059401	2.615	(0.148)	0.984	(0.027)	0.206	(0.011)			92M11	9
M059501	1.842	(0.203)	1.911	(0.095)	0.394	(0.011)			92M11	10
M060701	0.500	(0.028)	1.714	(0.055)			1.079	(0.073)	92M13	9
							0.134	(0.102)		
							-0.750	(0.177)		
							-0.462	(0.273)		
M061501	2.255	(0.118)	1.456	(0.034)	0.107	(0.008)			92M15	8
M061701	1.697	(0.118)	1.112	(0.033)	0.096	(0.010)			92M15	10

BEST COPY AVAILABLE

Table E-26 (continued)
IRT Parameters for Mathematics Main Samples
Algebra and Functions, Age 17/Grade 12

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
M061801	0.621	(0.035)	2.191	(0.070)			0.525	(0.060)	92M15	11
							-0.114	(0.117)		
							-1.511	(0.371)		
							1.099	(0.443)		
M062001	1.142	(0.082)	-0.932	(0.086)	0.205	(0.043)			92M10	6
M062201	1.474	(0.125)	1.272	(0.043)	0.087	(0.010)			92M10	8
M062301	1.895	(0.123)	1.498	(0.043)	0.098	(0.008)			92M10	9

Table E-27
IRT Parameters for Mathematics Main Samples Estimation, Age 17/Grade 12

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
M032001	0.602	(0.042)	-0.610	(0.137)	0.241	(0.044)			92M16	1
M032201	0.691	(0.047)	-2.102	(0.173)	0.305	(0.063)			92M16	3
M032301	0.810	(0.050)	-0.942	(0.109)	0.216	(0.046)			92M16	4
M032401	0.509	(0.053)	-0.150	(0.207)	0.351	(0.050)			92M16	5
M032501	0.388	(0.031)	-0.762	(0.200)	0.276	(0.044)			92M16	6
M032601	0.802	(0.053)	-0.883	(0.120)	0.293	(0.047)			92M16	7
M032701	0.626	(0.044)	-1.801	(0.190)	0.315	(0.064)			92M16	8
M032801	1.216	(0.080)	0.250	(0.045)	0.274	(0.019)			92M16	9
M032901	0.830	(0.056)	-0.395	(0.094)	0.267	(0.036)			92M16	10
M033001	0.573	(0.043)	0.484	(0.086)	0.123	(0.028)			92M16	11
M033101	0.582	(0.040)	-0.565	(0.133)	0.222	(0.043)			92M16	12
M033201	0.650	(0.058)	-0.871	(0.205)	0.399	(0.059)			92M16	13
M033301	1.059	(0.055)	-0.042	(0.044)	0.122	(0.020)			92M16	14
M033401	0.430	(0.053)	0.548	(0.206)	0.351	(0.043)			92M16	15
M033501	0.660	(0.055)	-0.261	(0.138)	0.308	(0.043)			92M16	16
M033601	0.715	(0.071)	1.291	(0.071)	0.189	(0.019)			92M16	17
M033701	0.952	(0.071)	0.503	(0.055)	0.237	(0.021)			92M16	18
M033801	1.191	(0.068)	-0.176	(0.049)	0.225	(0.023)			92M16	19
M033901	0.551	(0.046)	0.178	(0.123)	0.181	(0.038)			92M16	20
M034001	0.848	(0.086)	1.386	(0.066)	0.228	(0.016)			92M16	21
M034101	1.127	(0.070)	0.718	(0.035)	0.132	(0.013)			92M16	22

Table E-28
IRT Parameters for Writing Long-term Trend Samples
Grade 4

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
N000602	0.340	(0.017)	0.608	(0.067)			-2.651	(0.197)	84E	11
							2.651	(0.206)		
N000902	0.480	(0.033)	3.059	(0.075)			2.536	(0.085)	84G	7
							1.317	(0.124)		
							-3.853	(1.052)		
N001002	0.391	(0.024)	3.403	(0.082)			4.576	(0.120)	84G	8
							1.231	(0.143)		
							-5.806	(1.904)		
N014702	0.509	(0.032)	-0.472	(0.054)			0.868	(0.095)	84C	23
							-0.868	(0.079)		
N014802	0.716	(0.039)	1.708	(0.042)			2.017	(0.061)	84E	10
							0.585	(0.072)		
							-2.602	(0.350)		

Table E-29
IRT Parameters for Writing Long-term Trend Samples
Grade 8

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
N000302	0.649	(0.039)	2.034	(0.060)			1.831	(0.060)	84C	23
							0.301	(0.101)		
							-2.132	(0.431)		
N000402	0.384	(0.016)	2.372	(0.078)			4.289	(0.100)	84D	25
							-0.367	(0.137)		
							-3.921	(0.774)		
N000502	0.710	(0.027)	1.293	(0.045)			2.819	(0.072)	84E	10
							-0.404	(0.084)		
							-2.415	(0.379)		
N000602	0.370	(0.018)	-1.233	(0.073)			-2.235	(0.204)	84E	11
							2.235	(0.183)		
N000902	0.370	(0.022)	2.041	(0.063)			2.713	(0.120)	84G	7
							1.251	(0.128)		
							-3.965	(0.556)		
N001002	0.391	(0.019)	0.901	(0.071)			3.694	(0.175)	84G	8
							0.932	(0.108)		
							-4.626	(0.418)		

Table E-30
IRT Parameters for Writing Long-term Trend Samples
Grade 11

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
N000302	0.577	(0.029)	1.569	(0.050)			2.567	(0.069)	84C	23
							0.305	(0.087)		
							-2.872	(0.417)		
N000402	0.634	(0.024)	1.069	(0.054)			3.312	(0.079)	84D	25
							-0.502	(0.081)		
							-2.811	(0.351)		
N001002	0.285	(0.015)	1.078	(0.100)			4.413	(0.277)	84G	8
							1.791	(0.165)		
							-6.204	(0.614)		
N018002	0.524	(0.023)	1.054	(0.051)			2.827	(0.113)	84E	10
							0.130	(0.098)		
							-2.957	(0.379)		
N019002	0.291	(0.012)	1.226	(0.089)			1.687	(0.251)	84E	11
							4.240	(0.199)		
							-5.926	(0.505)		
N021002	0.767	(0.039)	1.143	(0.036)			1.712	(0.065)	84G	7
							0.328	(0.077)		
							-2.040	(0.297)		

Table E-31
IRT Parameters for Writing Main Samples
Age 9/Grade 4

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
W001001	0.478	(0.043)	0.565	(0.113)			2.536	(0.503)	92W3	1
							2.813	(0.307)		
							-0.231	(0.214)		
							-2.255	(0.410)		
							-2.862	(0.938)		
W001101	0.442	(0.040)	0.498	(0.114)			2.723	(0.515)	92W5	1
							2.637	(0.304)		
							-0.694	(0.247)		
							-1.606	(0.396)		
							-3.060	(0.877)		
W001201	0.694	(0.069)	0.936	(0.083)			3.368	(0.280)	92W6	1
							1.196	(0.149)		
							-0.154	(0.177)		
							-1.961	(0.415)		
							-2.449	(1.070)		
W001301	0.750	(0.075)	0.905	(0.075)			2.567	(0.234)	92W8	1
							1.627	(0.154)		
							-0.004	(0.153)		
							-1.834	(0.346)		
							-2.356	(0.869)		
W001401	0.443	(0.048)	1.400	(0.139)			3.544	(0.316)	92W9	1
							0.665	(0.235)		
							-1.312	(0.423)		
							-2.898	(1.253)		
W001501	0.520	(0.062)	1.007	(0.103)			2.047	(0.245)	92W11	1
							1.011	(0.213)		
							-0.747	(0.285)		
							-2.311	(0.665)		
W001601	0.484	(0.044)	0.912	(0.108)			2.046	(0.411)	92W4	1
							3.105	(0.302)		
							-0.171	(0.224)		
							-1.903	(0.444)		
							-3.078	(1.220)		
W001701	0.762	(0.079)	1.079	(0.078)			2.873	(0.208)	92W7	1
							1.153	(0.139)		
							-0.132	(0.176)		
							-2.072	(0.495)		
							-1.822	(1.026)		
W001801	0.483	(0.052)	1.331	(0.126)			3.495	(0.299)	92W10	1
							0.843	(0.210)		
							-1.119	(0.356)		
							-3.219	(1.229)		

Table E-32
IRT Parameters for Writing Main Samples
Age 13/Grade 8

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
W001601	0.494	(0.052)	0.250	(0.126)			1.547	(0.747)	92W4	1
							3.621	(0.523)		
							0.174	(0.229)		
							-1.621	(0.334)		
							-3.721	(0.905)		
W001701	0.550	(0.065)	0.426	(0.116)			3.782	(0.652)	92W7	1
							1.641	(0.264)		
							0.267	(0.218)		
							-2.144	(0.397)		
							-3.546	(1.150)		
W001801	0.554	(0.063)	0.217	(0.130)			3.508	(0.571)	92W10	1
							1.206	(0.220)		
							-1.264	(0.250)		
							-3.451	(0.701)		
W001901	0.567	(0.068)	0.072	(0.090)			2.916	(0.435)	92W13	1
							0.149	(0.254)		
							0.270	(0.255)		
							-1.262	(0.292)		
							-2.072	(0.461)		
W002001	0.415	(0.044)	1.266	(0.131)			0.534	(0.665)	92W3	1
							4.930	(0.593)		
							0.277	(0.296)		
							-1.202	(0.504)		
							-4.538	(2.040)		
W002101	0.396	(0.045)	1.567	(0.151)			3.711	(0.554)	92W5	1
							3.163	(0.342)		
							0.169	(0.328)		
							-1.844	(0.699)		
							-5.199	(3.614)		
W002201	0.600	(0.074)	1.249	(0.105)			3.836	(0.401)	92W6	1
							1.607	(0.210)		
							0.628	(0.224)		
							-1.585	(0.466)		
							-4.486	(2.692)		
W002301	0.426	(0.052)	0.713	(0.140)			4.814	(0.720)	92W8	1
							0.993	(0.291)		
							0.182	(0.304)		
							-2.966	(0.664)		
							-3.024	(1.480)		
W002401	0.508	(0.057)	1.255	(0.162)			4.166	(0.501)	92W9	1
							2.068	(0.226)		
							-1.865	(0.413)		
							-4.370	(2.139)		
W002501	0.576	(0.067)	1.412	(0.133)			3.584	(0.346)	92W11	1
							1.818	(0.196)		
							-1.323	(0.364)		
							-4.080	(2.015)		
W002601	0.648	(0.075)	-0.275	(0.099)			1.821	(0.612)	92W12	1
							2.098	(0.383)		
							0.425	(0.202)		
							-1.239	(0.211)		
							-3.105	(0.471)		

Table E-33
IRT Parameters for Writing Main Samples
Age 17/Grade 12

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
W002001	0.418	(0.046)	0.681	(0.141)			-0.381	(1.044)	92W3	1
							5.785	(0.948)		
							-0.276	(0.315)		
							-0.503	(0.413)		
							-4.625	(1.430)		
W002101	0.436	(0.047)	1.173	(0.150)			2.501	(0.571)	92W5	1
							3.625	(0.399)		
							-0.386	(0.319)		
							-1.609	(0.612)		
							-4.130	(2.251)		
W002201	0.592	(0.072)	0.174	(0.110)			2.442	(0.549)	92W6	1
							1.730	(0.317)		
							0.552	(0.222)		
							-1.681	(0.299)		
							-3.044	(0.692)		
W002301	0.446	(0.057)	0.287	(0.129)			3.733	(0.694)	92W8	1
							1.138	(0.326)		
							0.129	(0.303)		
							-1.732	(0.428)		
							-3.269	(0.952)		
W002401	0.547	(0.060)	0.552	(0.161)			3.712	(0.624)	92W9	1
							1.918	(0.244)		
							-1.804	(0.322)		
							-3.827	(1.217)		
W002501	0.610	(0.073)	0.608	(0.150)			3.698	(0.554)	92W11	1
							1.534	(0.216)		
							-1.665	(0.318)		
							-3.567	(1.172)		
W002601	0.709	(0.086)	-0.797	(0.110)			1.570	(0.994)	92W12	1
							2.039	(0.612)		
							0.842	(0.254)		
							-1.052	(0.179)		
							-3.399	(0.372)		
W002701	0.423	(0.053)	0.420	(0.132)			1.865	(0.635)	92W4	1
							2.615	(0.458)		
							0.468	(0.308)		
							-1.135	(0.382)		
							-3.814	(0.994)		
W002801	0.502	(0.062)	0.535	(0.122)			2.860	(0.549)	92W7	1
							1.933	(0.325)		
							0.545	(0.255)		
							-1.756	(0.392)		
							-3.581	(1.153)		
W002901	0.548	(0.066)	1.475	(0.148)			4.635	(0.532)	92W10	1
							2.069	(0.225)		
							-0.673	(0.332)		
							-2.659	(1.105)		
							-3.371	(3.788)		
W003001	0.666	(0.081)	0.012	(0.099)			3.132	(0.586)	92W13	1
							1.210	(0.228)		
							-0.684	(0.214)		
							-1.410	(0.303)		
							-2.248	(0.498)		

Table E-33 (continued)
IRT Parameters for Writing Main Samples
Age 17/Grade 12

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
W003101	0.754	(0.090)	0.830	(0.106)			3.543	(0.423)	92W14	1
							1.635	(0.182)		
							-0.351	(0.201)		
							-1.914	(0.475)		
							-2.912	(1.527)		

Table E-34
IRT Parameters for Science Long-term Trend Samples
Age 9

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
N400001	0.726	(0.061)	-1.057	(0.138)	0.495	(0.028)			86S1	6
N400101	1.011	(0.150)	1.932	(0.114)	0.551	(0.011)			86S1	15
N400102	0.698	(0.102)	1.694	(0.122)	0.475	(0.017)			86S1	16
N400301	0.823	(0.068)	0.068	(0.082)	0.482	(0.020)			86S1	8
N400401	0.738	(0.061)	-1.933	(0.185)	0.483	(0.036)			86S1	9
N400402	1.348	(0.079)	-1.088	(0.055)	0.292	(0.020)			86S1	10
N400403	0.426	(0.039)	-2.652	(0.388)	0.491	(0.058)			86S1	11
N400404	1.051	(0.066)	-0.841	(0.067)	0.365	(0.021)			86S1	12
N400405	0.593	(0.045)	-1.404	(0.161)	0.372	(0.033)			86S1	13
N400501	0.529	(0.056)	0.594	(0.125)	0.426	(0.024)			86S1	14
N400601	0.822	(0.057)	0.071	(0.066)	0.346	(0.020)			86S1	17
N400701	0.930	(0.056)	0.248	(0.047)	0.248	(0.017)			86S1	18
N400901	0.170	(0.028)	2.833	(0.452)	0.304	(0.028)			86S1	19
N401001	0.475	(0.040)	0.373	(0.104)	0.230	(0.024)			86S1	20
N401101	0.414	(0.086)	2.506	(0.282)	0.445	(0.023)			86S1	21
N401201	0.798	(0.125)	2.437	(0.147)	0.246	(0.012)			86S1	22
N401301	0.589	(0.057)	1.016	(0.095)	0.333	(0.020)			86S1	23
N401501	0.503	(0.087)	1.852	(0.182)	0.545	(0.020)			86S2	1
N401601	0.807	(0.064)	-0.621	(0.104)	0.553	(0.022)			86S2	2
N401702	0.336	(0.096)	3.296	(0.583)	0.624	(0.020)			86S2	4
N401703	0.294	(0.067)	2.323	(0.362)	0.547	(0.029)			86S2	5
N401801	1.040	(0.085)	0.140	(0.069)	0.578	(0.017)			86S2	6
N401802	1.384	(0.121)	-0.015	(0.060)	0.644	(0.015)			86S2	7
N401803	0.975	(0.093)	0.509	(0.080)	0.623	(0.016)			86S2	8
N401804	0.623	(0.087)	1.630	(0.127)	0.507	(0.017)			86S2	9
N401901	0.446	(0.071)	2.121	(0.183)	0.389	(0.021)			86S2	10
N402001	0.677	(0.043)	-1.558	(0.127)	0.313	(0.030)			86S2	11
N402002	0.673	(0.043)	-1.797	(0.140)	0.323	(0.032)			86S2	12
N402005	0.709	(0.059)	-0.061	(0.100)	0.510	(0.021)			86S2	15
N402101	0.690	(0.051)	0.075	(0.084)	0.400	(0.021)			86S2	16
N402201	0.246	(0.028)	0.505	(0.255)	0.298	(0.035)			86S2	17
N402401	0.373	(0.103)	4.100	(0.688)	0.374	(0.018)			86S2	18
N402501	0.920	(0.092)	1.826	(0.069)	0.232	(0.011)			86S2	19
N402602	0.401	(0.000)	-0.194	(0.194)	0.626	(0.023)			86S2	21
N402701	0.481	(0.060)	2.097	(0.141)	0.247	(0.018)			86S2	23
N402801	1.330	(0.083)	1.864	(0.050)	0.181	(0.007)			86S2	24
N402901	0.065	(0.014)	13.924	(2.934)	0.077	(0.016)			86S2	25
N403001	0.202	(0.033)	******	(1.712)	0.371	(0.074)			86S3	12
N403101	0.335	(0.030)	-6.720	(0.548)	0.330	(0.069)			86S3	13
N403201	0.351	(0.022)	-3.576	(0.311)	0.268	(0.058)			86S3	14
N403202	0.249	(0.020)	-1.949	(0.320)	0.251	(0.044)			86S3	15
N403301	0.483	(0.041)	-1.282	(0.211)	0.385	(0.039)			86S3	16
N403401	0.367	(0.048)	0.596	(0.217)	0.437	(0.033)			86S3	17
N403501	0.399	(0.043)	-0.167	(0.211)	0.407	(0.035)			86S3	18
N403502	0.413	(0.034)	-3.070	(0.395)	0.483	(0.062)			86S3	19
N403503	0.253	(0.032)	0.128	(0.306)	0.467	(0.033)			86S3	20
N403601	0.586	(0.045)	0.704	(0.078)	0.258	(0.020)			86S3	21
N403701	3.285	(0.000)	-0.381	(0.017)	0.330	(0.013)			86S3	22
N403702	3.146	(0.000)	-0.426	(0.021)	0.484	(0.014)			86S3	23
N403703	3.324	(0.208)	-0.331	(0.024)	0.371	(0.014)			86S3	24
N403801	0.394	(0.074)	2.200	(0.235)	0.473	(0.024)			86S3	25
N403803	0.365	(0.042)	-1.248	(0.367)	0.504	(0.047)			86S3	27
N403804	0.265	(0.029)	-0.744	(0.325)	0.434	(0.037)			86S3	28
N403901	0.511	(0.036)	-0.511	(0.116)	0.243	(0.028)			86S3	29
N404001	0.185	(0.024)	1.825	(0.297)	0.265	(0.028)			86S3	30
N404201	0.461	(0.047)	1.507	(0.106)	0.215	(0.021)			86S3	31

Table E-35
IRT Parameters for Science Long-term Trend Samples
Age 13

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
N400201	0.422	(0.034)	-1.683	(0.250)	0.279	(0.058)			86S1	16
N401201	0.934	(0.074)	0.191	(0.069)	0.308	(0.025)			86S1	28
N404501	0.920	(0.062)	-2.358	(0.135)	0.259	(0.057)			86S1	12
N404601	0.518	(0.057)	-0.012	(0.186)	0.312	(0.047)			86S1	13
N404701	0.539	(0.038)	-1.978	(0.201)	0.274	(0.058)			86S1	14
N404702	0.508	(0.055)	0.059	(0.178)	0.285	(0.045)			86S1	15
N404801	0.909	(0.077)	-1.927	(0.193)	0.516	(0.061)			86S1	20
N404802	1.838	(0.123)	-0.207	(0.042)	0.388	(0.020)			86S1	21
N404803	1.422	(0.103)	0.435	(0.039)	0.327	(0.017)			86S1	22
N404901	0.761	(0.055)	-0.519	(0.111)	0.281	(0.038)			86S1	17
N405001	1.084	(0.103)	0.733	(0.057)	0.383	(0.019)			86S1	23
N405101	0.976	(0.085)	0.910	(0.051)	0.239	(0.017)			86S1	24
N405201	1.042	(0.149)	1.233	(0.079)	0.500	(0.017)			86S1	25
N405301	1.060	(0.124)	1.344	(0.063)	0.309	(0.015)			86S1	26
N405401	1.044	(0.088)	1.053	(0.046)	0.198	(0.015)			86S1	27
N405501	1.151	(0.090)	0.392	(0.052)	0.339	(0.020)			86S1	29
N405601	0.433	(0.096)	2.003	(0.217)	0.341	(0.035)			86S1	30
N405701	1.312	(0.085)	0.554	(0.034)	0.201	(0.015)			86S1	31
N405801	1.270	(0.116)	1.175	(0.044)	0.241	(0.013)			86S1	32
N405901	1.192	(0.127)	1.751	(0.076)	0.218	(0.010)			86S1	33
N406001	1.454	(0.205)	2.490	(0.169)	0.153	(0.007)			86S1	34
N406101	1.335	(0.163)	2.261	(0.129)	0.187	(0.008)			86S1	35
N406201	0.978	(0.154)	2.348	(0.165)	0.111	(0.009)			86S1	36
N406301	0.298	(0.044)	0.385	(0.301)	0.584	(0.031)			86S2	10
N406302	0.386	(0.041)	-0.774	(0.265)	0.484	(0.042)			86S2	11
N406303	1.325	(0.164)	0.335	(0.079)	0.509	(0.026)			86S2	12
N406304	0.485	(0.094)	1.367	(0.184)	0.465	(0.033)			86S2	13
N406401	0.632	(0.083)	0.091	(0.188)	0.549	(0.036)			86S2	14
N406402	1.164	(0.088)	0.067	(0.061)	0.413	(0.022)			86S2	15
N406403	0.762	(0.072)	-1.380	(0.227)	0.572	(0.054)			86S2	16
N406404	1.115	(0.093)	-0.343	(0.091)	0.521	(0.027)			86S2	17
N406405	1.161	(0.099)	-0.182	(0.082)	0.534	(0.025)			86S2	18
N406501	0.812	(0.072)	0.729	(0.062)	0.242	(0.022)			86S2	19
N406601	0.423	(0.037)	-0.881	(0.227)	0.305	(0.050)			86S2	20
N406701	0.889	(0.081)	0.461	(0.072)	0.366	(0.024)			86S2	21
N406801	0.691	(0.055)	-1.911	(0.222)	0.474	(0.062)			86S2	22
N406802	0.491	(0.092)	1.010	(0.207)	0.523	(0.036)			86S2	23
N406803	0.674	(0.051)	-0.835	(0.152)	0.359	(0.045)			86S2	24
N406804	0.733	(0.052)	-1.204	(0.151)	0.379	(0.048)			86S2	25
N406805	1.535	(0.160)	1.183	(0.057)	0.573	(0.011)			86S2	26
N406806	0.343	(0.040)	0.258	(0.226)	0.412	(0.036)			86S2	27
N406901	0.696	(0.072)	0.356	(0.109)	0.366	(0.031)			86S2	28
N407001	0.468	(0.089)	1.394	(0.176)	0.402	(0.036)			86S2	29
N407101	1.202	(0.118)	1.790	(0.073)	0.171	(0.009)			86S2	30
N407201	1.362	(0.137)	0.953	(0.048)	0.439	(0.015)			86S2	31
N407301	0.410	(0.053)	0.872	(0.172)	0.238	(0.040)			86S2	32
N407302	1.115	(0.157)	1.606	(0.087)	0.435	(0.013)			86S2	33
N407601	0.844	(0.098)	1.366	(0.071)	0.261	(0.018)			86S2	35
N407701	0.690	(0.075)	1.248	(0.074)	0.202	(0.021)			86S2	37
N407801	0.885	(0.159)	2.072	(0.149)	0.273	(0.014)			86S2	38
N407901	0.503	(0.074)	1.167	(0.132)	0.288	(0.034)			86S2	39
N408001	1.387	(0.097)	0.887	(0.032)	0.202	(0.012)			86S2	34
N408201	1.203	(0.124)	1.913	(0.085)	0.179	(0.009)			86S2	40
N408301	1.613	(0.126)	0.806	(0.034)	0.342	(0.013)			86S3	10

Table E-35 (continued)
IRT Parameters for Science Long-term Trend Samples
Age 13

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
N408302	0.776	(0.059)	-1.290	(0.164)	0.440	(0.050)			86S3	11
N408303	0.738	(0.062)	-1.239	(0.190)	0.480	(0.052)			86S3	12
N408304	1.043	(0.077)	-1.061	(0.117)	0.471	(0.040)			86S3	13
N408401	0.326	(0.031)	-0.607	(0.237)	0.312	(0.042)			86S3	14
N408501	0.792	(0.061)	-0.451	(0.115)	0.339	(0.038)			86S3	15
N408502	0.573	(0.070)	1.181	(0.097)	0.233	(0.027)			86S3	16
N408601	0.555	(0.045)	-0.659	(0.170)	0.234	(0.050)			86S3	17
N408701	0.699	(0.109)	1.136	(0.112)	0.481	(0.025)			86S3	18
N408801	0.368	(0.052)	1.019	(0.207)	0.349	(0.038)			86S3	19
N408901	1.106	(0.115)	0.594	(0.069)	0.533	(0.019)			86S3	20
N408902	0.962	(0.077)	-1.663	(0.166)	0.517	(0.054)			86S3	21
N408903	1.028	(0.095)	0.613	(0.061)	0.406	(0.020)			86S3	22
N408904	0.715	(0.119)	1.303	(0.115)	0.498	(0.022)			86S3	23
N409001	0.645	(0.058)	0.054	(0.118)	0.278	(0.036)			86S3	24
N409101	0.816	(0.058)	-1.007	(0.132)	0.350	(0.046)			86S3	25
N409102	0.865	(0.078)	0.453	(0.073)	0.345	(0.024)			86S3	26
N409103	1.218	(0.135)	1.550	(0.070)	0.366	(0.012)			86S3	27
N409201	0.718	(0.094)	1.031	(0.094)	0.398	(0.024)			86S3	28
N409301	0.852	(0.058)	0.030	(0.068)	0.217	(0.026)			86S3	29
N409501	0.882	(0.100)	1.734	(0.084)	0.151	(0.013)			86S3	33
N409601	1.766	(0.117)	1.117	(0.034)	0.309	(0.010)			86S3	34
N409701	1.296	(0.114)	1.719	(0.064)	0.195	(0.009)			86S3	35

Table E-36
IRT Parameters for Science Long-term Trend Samples
Age 17

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
N400201	0.519	(0.035)	-2.520	(0.207)	0.245	(0.057)			86S1	12
N401201	0.901	(0.059)	-0.222	(0.077)	0.317	(0.028)			86S1	30
N404601	0.398	(0.032)	-1.375	(0.236)	0.279	(0.051)			86S1	13
N405001	0.390	(0.038)	-0.145	(0.211)	0.305	(0.043)			86S1	29
N405101	0.872	(0.058)	-0.020	(0.071)	0.286	(0.025)			86S3	14
N405201	0.824	(0.090)	0.706	(0.087)	0.463	(0.022)			86S1	31
N405401	0.876	(0.064)	0.685	(0.051)	0.208	(0.018)			86S3	19
N405501	0.630	(0.046)	-0.690	(0.139)	0.287	(0.042)			86S3	21
N406001	0.827	(0.114)	1.868	(0.102)	0.244	(0.014)			86S1	33
N406101	1.139	(0.103)	1.622	(0.057)	0.181	(0.010)			86S1	35
N406201	1.433	(0.089)	1.603	(0.042)	0.110	(0.007)			86S1	37
N406301	0.976	(0.112)	0.386	(0.102)	0.641	(0.020)			86S1	21
N406302	0.257	(0.028)	-1.748	(0.352)	0.421	(0.041)			86S1	22
N406303	0.697	(0.062)	-0.286	(0.141)	0.441	(0.036)			86S1	23
N406304	0.468	(0.048)	-0.493	(0.234)	0.441	(0.046)			86S1	24
N406401	0.887	(0.069)	-0.398	(0.105)	0.464	(0.031)			86S2	10
N406402	1.057	(0.072)	-0.625	(0.086)	0.435	(0.030)			86S2	11
N406403	1.003	(0.076)	-1.599	(0.141)	0.490	(0.048)			86S2	12
N406404	1.022	(0.066)	-1.217	(0.101)	0.403	(0.038)			86S2	13
N406405	0.986	(0.067)	-1.302	(0.115)	0.433	(0.041)			86S2	14
N406601	0.456	(0.040)	-0.975	(0.234)	0.328	(0.054)			86S1	28
N406801	0.616	(0.047)	-2.671	(0.233)	0.405	(0.063)			86S2	16
N406802	0.259	(0.035)	1.048	(0.266)	0.424	(0.030)			86S2	17
N406803	0.619	(0.047)	-1.408	(0.190)	0.403	(0.051)			86S2	18
N406804	0.592	(0.046)	-2.079	(0.232)	0.431	(0.060)			86S2	19
N406805	0.551	(0.060)	-0.015	(0.186)	0.449	(0.039)			86S2	20
N406806	0.316	(0.033)	-0.030	(0.230)	0.384	(0.036)			86S2	21
N406901	0.523	(0.049)	-0.369	(0.189)	0.305	(0.048)			86S2	27
N407001	0.317	(0.028)	-0.845	(0.235)	0.251	(0.043)			86S2	33
N407101	0.946	(0.084)	1.254	(0.050)	0.190	(0.014)			86S2	38
N407201	0.501	(0.047)	0.088	(0.154)	0.241	(0.041)			86S2	32
N407301	0.191	(0.024)	1.379	(0.288)	0.239	(0.029)			86S2	36
N407302	1.039	(0.129)	1.304	(0.066)	0.434	(0.015)			86S2	37
N407401	0.369	(0.037)	-0.747	(0.257)	0.463	(0.040)			86S2	28
N407403	0.625	(0.059)	-0.274	(0.162)	0.435	(0.039)			86S2	30
N407404	0.685	(0.051)	-1.776	(0.192)	0.438	(0.054)			86S2	31
N407701	0.585	(0.046)	0.814	(0.073)	0.138	(0.022)			86S2	35
N408101	1.177	(0.104)	1.458	(0.048)	0.191	(0.011)			86S1	38
N408301	0.776	(0.050)	-0.724	(0.104)	0.324	(0.035)			86S3	10
N408302	0.604	(0.045)	-2.305	(0.230)	0.422	(0.061)			86S3	11
N408303	0.707	(0.053)	-2.148	(0.203)	0.443	(0.059)			86S3	12
N408304	0.842	(0.060)	-2.099	(0.163)	0.422	(0.056)			86S3	13
N408601	0.411	(0.029)	-2.099	(0.225)	0.191	(0.054)			86S1	19
N408801	0.738	(0.054)	-0.260	(0.103)	0.299	(0.033)			86S3	24
N408901	0.803	(0.058)	-1.168	(0.141)	0.443	(0.043)			86S3	15
N408902	1.077	(0.086)	-2.044	(0.152)	0.500	(0.054)			86S3	16
N408903	0.616	(0.052)	-0.155	(0.132)	0.347	(0.035)			86S3	17
N408904	0.550	(0.052)	-0.293	(0.180)	0.396	(0.042)			86S3	18
N409301	0.691	(0.043)	-1.620	(0.146)	0.247	(0.052)			86S1	20
N409501	0.659	(0.058)	1.215	(0.065)	0.140	(0.018)			86S1	34
N409901	0.773	(0.048)	-1.079	(0.116)	0.277	(0.042)			86S1	18
N410003	0.262	(0.032)	-4.282	(0.624)	0.455	(0.061)			86S1	16
N410004	0.285	(0.030)	-2.127	(0.382)	0.471	(0.045)			86S1	17
N410101	0.738	(0.075)	-0.418	(0.170)	0.584	(0.035)			86S1	25

Table E-36 (continued)
IRT Parameters for Science Long-term Trend Samples
Age 17

NAEP ID	A	S.E.	B	S.E.	C	S.E.	D	S.E.	Block	Item
N410102	0.343	(0.037)	-0.768	(0.278)	0.497	(0.039)			86S1	26
N410103	0.367	(0.033)	-1.984	(0.305)	0.429	(0.050)			86S1	27
N410201	0.912	(0.096)	1.388	(0.060)	0.242	(0.015)			86S1	32
N410401	0.635	(0.103)	1.078	(0.136)	0.542	(0.026)			86S2	15
N410501	0.301	(0.023)	-1.073	(0.214)	0.156	(0.041)			86S2	22
N410601	2.038	(0.089)	1.222	(0.024)	0.140	(0.007)			86S2	23
N410602	0.475	(0.040)	-2.812	(0.316)	0.430	(0.066)			86S2	24
N410603	1.755	(0.118)	1.026	(0.034)	0.408	(0.011)			86S2	25
N410604	0.392	(0.035)	-2.961	(0.359)	0.418	(0.063)			86S2	26
N410701	0.908	(0.069)	0.750	(0.051)	0.227	(0.018)			86S2	34
N410801	0.737	(0.095)	1.466	(0.082)	0.293	(0.020)			86S2	39
N410901	1.011	(0.071)	1.183	(0.040)	0.114	(0.011)			86S2	40
N411001	1.208	(0.101)	1.326	(0.041)	0.171	(0.011)			86S2	41
N411101	0.583	(0.051)	0.118	(0.123)	0.248	(0.036)			86S3	22
N411201	0.948	(0.070)	0.448	(0.058)	0.300	(0.020)			86S3	23
N411301	1.268	(0.180)	2.708	(0.183)	0.148	(0.006)			86S3	20
N411401	2.237	(0.125)	0.475	(0.021)	0.226	(0.011)			86S3	25
N411501	1.102	(0.104)	1.457	(0.051)	0.194	(0.011)			86S3	26
N411502	0.734	(0.054)	-1.057	(0.151)	0.385	(0.046)			86S3	27
N411601	1.211	(0.084)	0.928	(0.035)	0.197	(0.013)			86S3	28
N411701	1.241	(0.088)	1.044	(0.035)	0.178	(0.012)			86S3	29
N411801	2.320	(0.124)	0.447	(0.020)	0.196	(0.010)			86S3	30
N411901	1.554	(0.100)	1.136	(0.032)	0.231	(0.010)			86S3	31
N412001	1.326	(0.111)	1.696	(0.061)	0.274	(0.009)			86S3	32

APPENDIX F

CONDITIONING VARIABLES AND CONTRAST CODINGS

APPENDIX F

Conditioning Variables and Contrast Codings

This appendix contains information about the conditioning variables used in the construction of plausible values for the 1992 National Assessment. Two kinds of conditioning variables were defined—continuous or quasi-continuous variables, such as school mathematics score or number of hours spent watching television, and categorical variables which made up the majority of the conditioning variables created from responses to student, teacher, and school demographic and background questionnaires.

Categorical conditioning variables derived from questionnaire or demographic variables were incorporated into the conditioning process by constructing a set of contrasts, each of which defines one or more of the variable's response options. A recoding procedure explodes the raw student responses into a binary series of one-degree-of-freedom "dummy" variables. Questionnaire or demographic variables that possess ordinal response options, such as number of hours spent watching television, were included in the conditioning process by creating linear and/or quadratic multi-degree-of-freedom contrasts. Continuous variables were included in the conditioning process in their original form.

This appendix gives the specifications used for constructing the conditioning variables. Table F-1 defines the information provided for each main sample variable. Conditioning variable data for the reading main samples are given in Table F-2; for mathematics main samples in Table F-3; and for writing main samples in Table F-4. Similar information for trend samples is given in Tables F-5, F-6, F-7, and F-8 respectively for reading, mathematics, writing, and science. Estimated effects for the trend conditioning variables are provided in Tables F-37 to F-48.

The linear conditioning model employed for the estimation of plausible values did not directly use the conditioning variable specifications listed in this appendix. To eliminate inherent instabilities in estimation encountered when using a large number of correlated variables, a principal component transformation of the correlation matrix obtained from the conditioning variable contrasts derived according to these primary specifications was performed. The principal components scores based on this transformation were used as the predictor variables in estimating the linear conditioning model. The estimated effects for the principal components for the reading main samples are given in Tables F-9 to F-16; for mathematics main samples in Tables F-17 to F-31; and for writing in Tables F-32 to F-36. The proportions of variance of the conditioning contrast acounted for the principal components are given in Tables F-49 to F-51 for reading, Tables F-52 to F-54 for mathematics, and Tables F-55 to F-57 for writing.

Table F-1
Description of Data Provided for Each Conditioning Variable

Title	Description
CONDITIONING ID	An unique eight-character ID assigned to identify each conditioning variable corresponding to a particular background or subject area question within the entire pool of conditioning variables. The first four characters identify the origin of the variable: BACK (background questionnaire), READ (student reading questionnaire), SCHL (school questionnaire), TCHR (background part of teacher questionnaire), and TRED (reading classroom part of teacher questionnaire). The second four digits represent the sequential position within each origin group.
DESCRIPTION	A short description of the conditioning variable.
GRADES/ASSESSMENTS	Three characters identifying assessment ("S" for state, "N" for national) and grade (04, 08, and 12) in which the conditioning variable was used.
GROUP LABEL	A descriptive eight-character label identifying the conditioning variable.
NAEP ID	The seven-character NAEP database identification for the conditioning variable.
TYPE OF CONTRAST	The type of conditioning variable. "CLASS" identifies a categorical conditioning variable and "SCALE" identifies continuous or quasi-continuous conditioning variables.
LENGTH OF CONTRAST FIELD	The number of columns (or length of the contrast field) for the conditioning variable within the entire conditioning variable vector. The length is associated with the number of explicit contrasts comprising categorical conditioning variables.
DEGREES OF FREEDOM	The number of degrees of freedom for each contrast constructed for the conditioning variable.
NUMBER OF SPECIFICATION RECORDS	The number of unique contrasts corresponding to each conditioning variable. For each contrast a specifications record is given with the following information: a sequential identification number, an eight-character descriptive label corresponding to the associated questionnaire option(s), a "collapsing code string" enclosed in parentheses specifying the database values to be merged to form the contrast, the contrast itself, and a short description of the contrast.

Table F-2
Conditioning Variables for Reading Main Samples

```
CONDITIONING ID:          BACK0001
DESCRIPTION:              GRAND MEAN
GRADES/ASSESSMENTS:       N04, S04, N08, S08, N12
GROUP LABEL:              OVERALL           LENGTH OF CONTRAST FIELD      :  1
NAEP ID:                  BKSER             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         SCALE             NUMBER OF SPECIFICATION RECORDS:  1
001   OVERALL    (@         )  1            GRAND MEAN

CONDITIONING ID:          BACK0002
DESCRIPTION:              GENDER (DERIVED)
GRADES/ASSESSMENTS:       N04, S04, N08, S08, N12
GROUP LABEL:              GENDER            LENGTH OF CONTRAST FIELD      :  1
NAEP ID:                  DSEX              DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  2
001   MALE       (1         )  0            GENDER:  MALE
002   FEMALE     (2         )  1            GENDER:  FEMALE

CONDITIONING ID:          BACK0003
DESCRIPTION:              ETHNICITY/RACE (DERIVED)
GRADES/ASSESSMENTS:       N04, S04, N08, S08, N12
GROUP LABEL:              ETHNICTY          LENGTH OF CONTRAST FIELD      :  3
NAEP ID:                  DRACE             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  4
001   WHIT/AOM   (1,5,6,M   )  000          ETHNICITY:  WHITE, AMERICAN INDIAN, UNCLASSIFIED, MISSING
002   BLACK      (2         )  100          ETHNICITY:  BLACK
003   HISPANIC   (3         )  010          ETHNICITY:  HISPANIC
004   ASIAN      (4         )  001          ETHNICITY:  ASIAN AMERICAN

CONDITIONING ID:          BACK0004
DESCRIPTION:              SIZE AND TYPE OF COMMUNITY
GRADES/ASSESSMENTS:       N04, N08, N12
GROUP LABEL:              STOC              LENGTH OF CONTRAST FIELD      :  2
NAEP ID:                  STOC              DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  3
001   LO_METRO   (2         )  00           STOC:  LOW METROPOLITAN
002   HI_METRO   (3         )  10           STOC:  HIGH METROPOLITAN
003   STOC-OTH   (1,4-7,M   )  01           STOC:  OTHER

CONDITIONING ID:          BACK0006
DESCRIPTION:              REGION OF THE COUNTRY
GRADES/ASSESSMENTS:       N04, N08, N12
GROUP LABEL:              REGION            LENGTH OF CONTRAST FIELD      :  3
NAEP ID:                  REGION            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  4
001   N_EAST     (1         )  000          REGION:  NORTHEAST
002   S_EAST     (2         )  100          REGION:  SOUTHEAST
003   CENTRAL    (3         )  010          REGION:  CENTRAL
004   WEST       (4         )  001          REGION:  WEST

CONDITIONING ID:          BACK0007
DESCRIPTION:              PARENTS' HIGHEST LEVEL OF EDUCATION
GRADES/ASSESSMENTS:       N04, S04, N08, S08, N12
GROUP LABEL:              PARED             LENGTH OF CONTRAST FIELD      :  4
NAEP ID:                  PARED             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  5
001   <HI_SCH    (1         )  0000         PARED:  LESS THAN HIGH SCHOOL
002   HS_GRAD    (2         )  1000         PARED:  HIGH SCHOOL GRADUATE
003   POST_HS    (3         )  0100         PARED:  POST HIGH SCHOOL
004   COL_GRAD   (4         )  0010         PARED:  COLLEGE GRADUATE
005   PARED-?    (M,IDK     )  0001         PARED:  MISSING, I DON'T KNOW

CONDITIONING ID:          BACK0008
DESCRIPTION:              ITEMS IN THE HOME (NEWSPAPER, > 25 BOOKS, ENCYCLOPEDIA, MAGAZINES)
GRADES/ASSESSMENTS:       N04, S04, N08, S08, N12
```

Table F-2 (continued)
Conditioning Variables for Reading Main Samples

```
GROUP LABEL:              HOMEITMS          LENGTH OF CONTRAST FIELD      :  2
NAEP ID:                  HOMEEN2           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  3
001   HITEM<=2   (1,M        )  00          ITEMS IN HOME:  ZERO TO TWO ITEMS, MISSING
002   HITEM=3    (2          )  10          ITEMS IN HOME:  THREE ITEMS
003   HITEM=4    (3          )  01          ITEMS IN HOME:  FOUR ITEMS

CONDITIONING ID:          BACK0009
DESCRIPTION:              HOURS OF TV WATCHING (LINEAR)
GRADES/ASSESSMENTS:       N04, S04, N08, S08, N12
GROUP LABEL:              TVWATCHL          LENGTH OF CONTRAST FIELD      :  1
NAEP ID:                  B001801           DEGREES OF FREEDOM PER CONTRAST:  6
TYPE OF CONTRAST:         SCALE             NUMBER OF SPECIFICATION RECORDS:  7
001   TV-LIN1    (1          )  0           TV WATCHING (LINEAR):  NONE
002   TV-LIN2    (2          )  1           TV WATCHING (LINEAR):  ONE HOUR OR LESS PER DAY
003   TV-LIN3    (3          )  2           TV WATCHING (LINEAR):  TWO HOURS PER DAY
004   TV-LIN4    (4,M        )  3           TV WATCHING (LINEAR):  THREE HOURS PER DAY
005   TV-LIN5    (5          )  4           TV WATCHING (LINEAR):  FOUR HOURS PER DAY
006   TV-LIN6    (6          )  5           TV WATCHING (LINEAR):  FIVE HOURS PER DAY
007   TV-LIN7    (7          )  6           TV WATCHING (LINEAR):  SIX OR MORE HOURS PER DAY

CONDITIONING ID:          BACK0010
DESCRIPTION:              HOURS OF TV WATCHING (QUADRATIC)
GRADES/ASSESSMENTS:       N04, S04, N08, S08, N12
GROUP LABEL:              TVWATCHQ          LENGTH OF CONTRAST FIELD      :  2
NAEP ID:                  B001801           DEGREES OF FREEDOM PER CONTRAST:  6
TYPE OF CONTRAST:         SCALE             NUMBER OF SPECIFICATION RECORDS:  7
001   TV-QUAD1   (1          )  00          TV WATCHING (QUADRATIC):  NONE
002   TV-QUAD2   (2          )  01          TV WATCHING (QUADRATIC):  ONE HOUR OR LESS PER DAY
003   TV-QUAD3   (3          )  04          TV WATCHING (QUADRATIC):  TWO HOURS PER DAY
004   TV-QUAD4   (4,M        )  09          TV WATCHING (QUADRATIC):  THREE HOURS PER DAY
005   TV-QUAD5   (5          )  16          TV WATCHING (QUADRATIC):  FOUR HOURS PER DAY
006   TV-QUAD6   (6          )  25          TV WATCHING (QUADRATIC):  FIVE HOURS PER DAY
007   TV-QUAD7   (7          )  36          TV WATCHING (QUADRATIC):  SIX OR MORE HOURS PER DAY

CONDITIONING ID:          BACK0011
DESCRIPTION:              HOME LANGUAGE MINORITY (HOW OFTEN PEOPLE IN HOME SPEAK LANG OTHER THE ENGLISH?)
GRADES/ASSESSMENTS:       N04, S04, N08, S08, N12
GROUP LABEL:              HOMELANG          LENGTH OF CONTRAST FIELD      :  1
NAEP ID:                  B003201           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  2
001   HL-NEV/?   (1,M        )  0           HOME LANGUAGE MINORITY:  NEVER, MISSING
002   HL-SM/AL   (2,3        )  1           HOME LANGUAGE MINORITY:  SOMTIMES, ALWAYS

CONDITIONING ID:          BACK0012
DESCRIPTION:              HOMEWORK ASSIGNED? (GRADE 4)
GRADES/ASSESSMENTS:       N04, S04
GROUP LABEL:              HW-CORE4          LENGTH OF CONTRAST FIELD      :  2
NAEP ID:                  B006601           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  3
001   HW4-MISS   (M          )  00          HOMEWORK ASSIGNED?:  MISSING
002   HW4-NONE   (1          )  10          HOMEWORK ASSIGNED?:  NO HOMEWORK ASSIGNED
003   HW4-YES    (2-5        )  01          HOMEWORK ASSIGNED?:  YES

CONDITIONING ID:          BACK0013
DESCRIPTION:              AMOUNT OF HOMEWORK (LINEAR) (GRADE 4)
GRADES/ASSESSMENTS:       N04, S04
GROUP LABEL:              HMWRKL4           LENGTH OF CONTRAST FIELD      :  1
NAEP ID:                  B006601           DEGREES OF FREEDOM PER CONTRAST:  3
TYPE OF CONTRAST:         SCALE             NUMBER OF SPECIFICATION RECORDS:  4
001   HW4-LIN1   (1,2,M      )  0           AMOUNT OF HOMEWORK (LINEAR):  DON'T HAVE, DON'T DO, MISSING
002   HW4-LIN2   (3          )  1           AMOUNT OF HOMEWORK (LINEAR):  ONE HALF HOUR
003   HW4-LIN3   (4          )  2           AMOUNT OF HOMEWORK (LINEAR):  ONE HOUR
004   HW4-LIN4   (5          )  3           AMOUNT OF HOMEWORK (LINEAR):  MORE THAN ONE HOUR
```

Table F-2 (continued)
Conditioning Variables for Reading Main Samples

```
CONDITIONING ID:          BACK0014
DESCRIPTION:              AMOUNT OF HOMEWORK (QUADRATIC) (GRADE 4)
GRADES/ASSESSMENTS:       N04, S04
GROUP LABEL:              HMWRKQ4             LENGTH OF CONTRAST FIELD      :   1
NAEP ID:                  B006601             DEGREES OF FREEDOM PER CONTRAST:   3
TYPE OF CONTRAST:         SCALE               NUMBER OF SPECIFICATION RECORDS:   4
001   HW4QUAD1   (1,2,M      )  0             AMOUNT OF HOMEWORK (QUAD):  DON'T HAVE, DON'T ANY, MISSING
002   HW4QUAD2   (3          )  1             AMOUNT OF HOMEWORK (QUADRATIC):  ONE HALF HOUR
003   HW4QUAD3   (4          )  4             AMOUNT OF HOMEWORK (QUADRATIC):  ONE HOUR
004   HW4QUAD4   (5          )  9             AMOUNT OF HOMEWORK (QUADRATIC):  MORE THAN ONE HOUR

CONDITIONING ID:          BACK0015
DESCRIPTION:              HOMEWORK ASSIGNED?
GRADES/ASSESSMENTS:       N08, S08, N12
GROUP LABEL:              HW-CORE             LENGTH OF CONTRAST FIELD      :   2
NAEP ID:                  B003901             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:         CLASS               NUMBER OF SPECIFICATION RECORDS:   3
001   HWC-MISS   (M          )  00            HOMEWORK ASSIGNED?:  MISSING
002   HWC-NONE   (1          )  10            HOMEWORK ASSIGNED?:  NO HOMEWORK ASSIGNED
003   HWC-YES    (2-6        )  01            HONEWORK ASSIGNED?:  YES

CONDITIONING ID:          BACK0016
DESCRIPTION:              AMOUNT OF HOMEWORK (LINEAR)
GRADES/ASSESSMENTS:       N08, S08, N12
GROUP LABEL:              HMWRKL              LENGTH OF CONTRAST FIELD      :   1
NAEP ID:                  B003901             DEGREES OF FREEDOM PER CONTRAST:   4
TYPE OF CONTRAST:         SCALE               NUMBER OF SPECIFICATION RECORDS:   5
001   HW-LIN1    (1,2,M      )  0             AMOUNT OF HOMEWORK (LINEAR):  DON'T HAVE, DON'T DO, MISSING
002   HW-LIN2    (3          )  1             AMOUNT OF HOMEWORK (LINEAR):  ONE HALF HOUR
003   HW-LIN3    (4          )  2             AMOUNT OF HOMEWORK (LINEAR):  ONE HOUR
004   HW-LIN4    (5          )  3             AMOUNT OF HOMEWORK (LINEAR):  TWO HOURS
005   HW-LIN5    (6          )  4             AMOUNT OF HOMEWORK (LINEAR):  MORE THAN TWO HOURS

CONDITIONING ID:          BACK0017
DESCRIPTION:              AMOUNT OF HOMEWORK (QUADRATIC)
GRADES/ASSESSMENTS:       N08, S08, N12
GROUP LABEL:              HMWRKQ              LENGTH OF CONTRAST FIELD      :   2
NAEP ID:                  B003901             DEGREES OF FREEDOM PER CONTRAST:   4
TYPE OF CONTRAST:         SCALE               NUMBER OF SPECIFICATION RECORDS:   5
001   HW-QUAD1   (1,2,M      )  00            AMOUNT OF HOMEWORK (QUAD):  DON'T HAVE, DON'T DO, MISSING
002   HW-QUAD2   (3          )  01            AMOUNT OF HOMEWORK (QUADRATIC):  ONE HALF HOUR
003   HW-QUAD3   (4          )  04            AMOUNT OF HOMEWORK (QUADRATIC):  ONE HOUR
004   HW-QUAD4   (5          )  09            AMOUNT OF HOMEWORK (QUADRATIC):  TWO HOURS
005   HW-QUAD5   (6          )  16            AMOUNT OF HOMEWORK (QUADRATIC):  MORE THAN TWO HOURS

CONDITIONING ID:          BACK0018
DESCRIPTION:              PERCENT WHITE STUDENTS IN SCHOOL
GRADES/ASSESSMENTS:       N04, S04, N08, S08, N12
GROUP LABEL:              %WHITE              LENGTH OF CONTRAST FIELD      :   2
NAEP ID:                  PCTWHT              DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:         CLASS               NUMBER OF SPECIFICATION RECORDS:   3
001   PREDOM/?   (80-110,M   )  00            PREDOMINANTLY WHITE, MISSING
002   MINORITY   (0-49       )  10            WHITE MINORITY
003   INTEGRAT   (50-79      )  01            INTEGRATED

CONDITIONING ID:          BACK0019
DESCRIPTION:              MODAL AGE, MODAL GRADE (DERIVED)
GRADES/ASSESSMENTS:       N04, N08, N12
GROUP LABEL:              AGE/GRAD            LENGTH OF CONTRAST FIELD      :   4
NAEP ID:                  MODGRAG             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:         CLASS               NUMBER OF SPECIFICATION RECORDS:   5
001   <MA/=MG    (1          )  0000          LESS THAN MODAL AGE, MODAL GRADE
002   =MA/<MG    (2          )  1000          MODAL AGE, LESS THAN MODAL GRADE
003   =MA/=MG    (3          )  0100          MODAL AGE, MODAL GRADE, MISSING
004   =MA/>MG    (4          )  0010          MODAL AGE, GREATER THAN MODAL GRADE
```

Table F-2 (continued)
Conditioning Variables for Reading Main Samples

```
005   >MA/=MG   (5            )   0001         GREATER THAN MODAL AGE, MODAL GRADE

CONDITIONING ID:           BACK0020
DESCRIPTION:               SCHOOL TYPE:  PUBLIC/NON-PUBLIC
GRADES/ASSESSMENTS:        N04, N08, N12
GROUP LABEL:               SCH_TYPE            LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                   SCHTYPE             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS               NUMBER OF SPECIFICATION RECORDS:   2
001   PUBLIC    (1            )   0            PUBLIC SCHOOL
002   NON_PUBL  (2-5,M        )   1            PRIVATE, CATHOLIC, BIA, DEPARTMENT OF DEFENSE, MISSING

CONDITIONING ID:           BACK0021
DESCRIPTION:               SINGLE/MULTIPLE PARENT(S) AT HOME
GRADES/ASSESSMENTS:        N04, S04, N08, S08, N12
GROUP LABEL:               PARENTS             LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                   SINGLEP             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS               NUMBER OF SPECIFICATION RECORDS:   2
001   NOT2PARS  (2-4,M        )   0            NOT TWO PARENTS, MISSING
002   2PARENTS  (1            )   1            BOTH FATHER AND MOTHER AT HOME

CONDITIONING ID:           BACK0022
DESCRIPTION:               MOTHER AT HOME
GRADES/ASSESSMENTS:        N04, S04, N08, S08, N12
GROUP LABEL:               MOM@HOME            LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                   B005601             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS               NUMBER OF SPECIFICATION RECORDS:   2
001   MOM@HM-N  (2,M          )   0            MOTHER AT HOME:  NO, MISSING
002   MOM@HM-Y  (1            )   1            MOTHER AT HOME:  YES

CONDITIONING ID:           BACK0023
DESCRIPTION:               PAGES READ FOR SCHOOL AND HOMEWORK EACH DAY (CONTRAST 1)
GRADES/ASSESSMENTS:        N04, S04, N08, S08, N12
GROUP LABEL:               PGSREAD1            LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                   B001101             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS               NUMBER OF SPECIFICATION RECORDS:   2
001   <=5_PGS   (5,M          )   0            PAGES READ (1):  5 OR FEWER PAGES, MISSING
002   >=6_PGS   (1-4          )   1            PAGES READ (1):  MORE THAN 20, 16-20, 11-15, 6-10

CONDITIONING ID:           BACK0024
DESCRIPTION:               PAGES READ FOR SCHOOL AND HOMEWORK EACH DAY (CONTRAST 2)
GRADES/ASSESSMENTS:        N04, S04, N08, S08, N12
GROUP LABEL:               PGSREAD2            LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                   B001101             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS               NUMBER OF SPECIFICATION RECORDS:   2
001   <=10_PGS  (4,5,M        )   0            PAGES READ (2):  6-10 PAGES, 5 OR FEWER PAGES, MISSING
002   >=11_PGS  (1-3          )   1            PAGES READ (2):  MORE THAN 20 PAGES, 16-20, 11-15 PAGES

CONDITIONING ID:           BACK0025
DESCRIPTION:               WENT TO PRESCHOOL?
GRADES/ASSESSMENTS:        N04, S04
GROUP LABEL:               PRESCH              LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                   B004201             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS               NUMBER OF SPECIFICATION RECORDS:   2
001   PRESCH-N  (2,3,IDK,M    )   0            WENT TO PRESCHOOL?:  NO, I DON'T KNOW, MISSING
002   PRESCH-Y  (1            )   1            WENT TO PRESCHOOL?:  YES

CONDITIONING ID:           BACK0026
DESCRIPTION:               DAYS OF SCHOOL MISSED LAST MONTH
GRADES/ASSESSMENTS:        N08, S08, N12
GROUP LABEL:               SCH_MISS            LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                   S004001             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS               NUMBER OF SPECIFICATION RECORDS:   2
001   >=3_DAYS  (3-5,M        )   0            DAYS SCHOOL MISSED:  3 OR 4, 5 TO 10, MORE THAN 10, MISSING
002   <=2_DAYS  (1,2          )   1            DAYS SCHOOL MISSED:  NONE, 1 OR 2 DAYS
```

Table F-2 (continued)
Conditioning Variables for Reading Main Samples

```
CONDITIONING ID:          BACK0027
DESCRIPTION:              HIGH SCHOOL PROGRAM
GRADES/ASSESSMENTS:       N12
GROUP LABEL:              HS_PROG          LENGTH OF CONTRAST FIELD      :  2
NAEP ID:                  B005001          DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS            NUMBER OF SPECIFICATION RECORDS:  3
001   GENERL/?   (1,M       )  00          HS PROGRAM:  GENERAL, MISSING
002   COL_PREP   (2         )  10          HS PROGRAM:  COLLEGE PREPARATORY
003   VOC/TECH   (3         )  01          HS PROGRAM:  VOCATIONAL, TECHNICAL

CONDITIONING ID:          BACK0028
DESCRIPTION:              NUMBER OF SEMESTERS ENGLISH/LITERATURE/WRITING (MISSING/NOT-MISSING)
GRADES/ASSESSMENTS:       N12
GROUP LABEL:              #SEM_ENG         LENGTH OF CONTRAST FIELD      :  1
NAEP ID:                  B007101          DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS            NUMBER OF SPECIFICATION RECORDS:  2
001   #ENG-?     (M         )  0           NUMBER OF SEMESTERS ENGLISH MISSING
002   #ENG-Y     (1-9       )  1           NUMBER OF SEMESTERS ENGLISH NOT-MISSING

CONDITIONING ID:          BACK0029
DESCRIPTION:              NUMBER OF SEMESTERS ENGLISH/LITERATURE/WRITING (LINEAR)
GRADES/ASSESSMENTS:       N12
GROUP LABEL:              #ENG-LIN         LENGTH OF CONTRAST FIELD      :  1
NAEP ID:                  B007101          DEGREES OF FREEDOM PER CONTRAST:  8
TYPE OF CONTRAST:         SCALE            NUMBER OF SPECIFICATION RECORDS: 10
001   #ENG-L01   (1         )  1           NUMBER OF SEMESTERS ENGLISH LINEAR
002   #ENG-L02   (2         )  2           NUMBER OF SEMESTERS ENGLISH LINEAR
003   #ENG-L03   (3         )  3           NUMBER OF SEMESTERS ENGLISH LINEAR
004   #ENG-L04   (4         )  4           NUMBER OF SEMESTERS ENGLISH LINEAR
005   #ENG-L05   (5         )  5           NUMBER OF SEMESTERS ENGLISH LINEAR
006   #ENG-L06   (6         )  6           NUMBER OF SEMESTERS ENGLISH LINEAR
007   #ENG-L07   (7         )  7           NUMBER OF SEMESTERS ENGLISH LINEAR
008   #ENG-L08   (8         )  8           NUMBER OF SEMESTERS ENGLISH LINEAR
009   #ENG-L09   (9         )  9           NUMBER OF SEMESTERS ENGLISH LINEAR
010   #ENG-L10   (M         )  0           NUMBER OF SEMESTERS ENGLISH LINEAR (MISSING)

CONDITIONING ID:          BACK0030
DESCRIPTION:              NUMBER OF SEMESTERS MATHEMATICS (MISSING/NOT-MISSING)
GRADES/ASSESSMENTS:       N12
GROUP LABEL:              #SEM_MAT         LENGTH OF CONTRAST FIELD      :  1
NAEP ID:                  B007102          DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS            NUMBER OF SPECIFICATION RECORDS:  2
001   #MATH-?    (M         )  0           NUMBER OF SEMESTERS MATHEMATICS MISSING
002   #MATH-Y    (1-9       )  1           NUMBER OF SEMESTERS MATHEMATICS NOT-MISSING

CONDITIONING ID:          BACK0031
DESCRIPTION:              NUMBER OF SEMESTERS MATHEMATICS/LITERATURE/WRITING (LINEAR)
GRADES/ASSESSMENTS:       N12
GROUP LABEL:              #MAT-LIN         LENGTH OF CONTRAST FIELD      :  1
NAEP ID:                  B007102          DEGREES OF FREEDOM PER CONTRAST:  8
TYPE OF CONTRAST:         SCALE            NUMBER OF SPECIFICATION RECORDS: 10
001   #MAT-L01   (1         )  1           NUMBER OF SEMESTERS MATHEMATICS LINEAR
002   #MAT-L02   (2         )  2           NUMBER OF SEMESTERS MATHEMATICS LINEAR
003   #MAT-L03   (3         )  3           NUMBER OF SEMESTERS MATHEMATICS LINEAR
004   #MAT-L04   (4         )  4           NUMBER OF SEMESTERS MATHEMATICS LINEAR
005   #MAT-L05   (5         )  5           NUMBER OF SEMESTERS MATHEMATICS LINEAR
006   #MAT-L06   (6         )  6           NUMBER OF SEMESTERS MATHEMATICS LINEAR
007   #MAT-L07   (7         )  7           NUMBER OF SEMESTERS MATHEMATICS LINEAR
008   #MAT-L08   (8         )  8           NUMBER OF SEMESTERS MATHEMATICS LINEAR
009   #MAT-L09   (9         )  9           NUMBER OF SEMESTERS MATHEMATICS LINEAR
010   #MAT-L10   (M         )  0           NUMBER OF SEMESTERS MATHEMATICS LINEAR (MISSING)

CONDITIONING ID:          BACK0032
DESCRIPTION:              NUMBER OF SEMESTERS SCIENCE (MISSING/NOT-MISSING)
GRADES/ASSESSMENTS:       N12
```

BEST COPY AVAILABLE

Table F-2 (continued)
Conditioning Variables for Reading Main Samples

```
GROUP LABEL:                #SEM_SCI          LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                    B007103           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS             NUMBER OF SPECIFICATION RECORDS:  2
001   #SCI-?     (M            )  0           NUMBER OF SEMESTERS SCIENCE MISSING
002   #SCI-Y     (1-9          )  1           NUMBER OF SEMESTERS SCIENCE NOT-MISSING

CONDITIONING ID:            BACK0033
DESCRIPTION:                NUMBER OF SEMESTERS SCIENCE/LITERATURE/WRITING (LINEAR)
GRADES/ASSESSMENTS:         N12
GROUP LABEL:                #SCI-LIN          LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                    B007103           DEGREES OF FREEDOM PER CONTRAST:  8
TYPE OF CONTRAST:           SCALE             NUMBER OF SPECIFICATION RECORDS: 10
001   #SCI-L01   (1            )  1           NUMBER OF SEMESTERS SCIENCE LINEAR
002   #SCI-L02   (2            )  2           NUMBER OF SEMESTERS SCIENCE LINEAR
003   #SCI-L03   (3            )  3           NUMBER OF SEMESTERS SCIENCE LINEAR
004   #SCI-L04   (4            )  4           NUMBER OF SEMESTERS SCIENCE LINEAR
005   #SCI-L05   (5            )  5           NUMBER OF SEMESTERS SCIENCE LINEAR
006   #SCI-L06   (6            )  6           NUMBER OF SEMESTERS SCIENCE LINEAR
007   #SCI-L07   (7            )  7           NUMBER OF SEMESTERS SCIENCE LINEAR
008   #SCI-L08   (8            )  8           NUMBER OF SEMESTERS SCIENCE LINEAR
009   #SCI-L09   (9            )  9           NUMBER OF SEMESTERS SCIENCE LINEAR
010   #SCI-L10   (M            )  0           NUMBER OF SEMESTERS SCIENCE LINEAR (MISSING)

CONDITIONING ID:            BACK0034
DESCRIPTION:                NUMBER OF SEMESTERS HISTORY/SOCIAL STUDIES/GEOGRAPHY (MISSING/NOT-MISSING)
GRADES/ASSESSMENTS:         N12
GROUP LABEL:                #SEM_HIS          LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                    B007104           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS             NUMBER OF SPECIFICATION RECORDS:  2
001   #HIS-?     (M            )  0           NUMBER OF SEMESTERS HISTORY MISSING
002   #HIS-Y     (1-9          )  1           NUMBER OF SEMESTERS HISTORY NOT-MISSING

CONDITIONING ID:            BACK0035
DESCRIPTION:                NUMBER OF SEMESTERS HISTORY/SOCIAL STUDIES/GEOGRAPHY (LINEAR)
GRADES/ASSESSMENTS:         N12
GROUP LABEL:                #HIS-LIN          LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                    B007104           DEGREES OF FREEDOM PER CONTRAST:  8
TYPE OF CONTRAST:           SCALE             NUMBER OF SPECIFICATION RECORDS: 10
001   #HIS-L01   (1            )  1           NUMBER OF SEMESTERS HISTORY LINEAR
002   #HIS-L02   (2            )  2           NUMBER OF SEMESTERS HISTORY LINEAR
003   #HIS-L03   (3            )  3           NUMBER OF SEMESTERS HISTORY LINEAR
004   #HIS-L04   (4            )  4           NUMBER OF SEMESTERS HISTORY LINEAR
005   #HIS-L05   (5            )  5           NUMBER OF SEMESTERS HISTORY LINEAR
006   #HIS-L06   (6            )  6           NUMBER OF SEMESTERS HISTORY LINEAR
007   #HIS-L07   (7            )  7           NUMBER OF SEMESTERS HISTORY LINEAR
008   #HIS-L08   (8            )  8           NUMBER OF SEMESTERS HISTORY LINEAR
009   #HIS-L09   (9            )  9           NUMBER OF SEMESTERS HISTORY LINEAR
010   #HIS-L10   (M            )  0           NUMBER OF SEMESTERS HISTORY LINEAR (MISSING)

CONDITIONING ID:            BACK0036
DESCRIPTION:                NUMBER OF SEMESTERS FOREIGN LANGUAGE(S) (MISSING/NOT-MISSING)
GRADES/ASSESSMENTS:         N12
GROUP LABEL:                #SEM_LAN          LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                    B007105           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS             NUMBER OF SPECIFICATION RECORDS:  2
001   #FLANG-?   (M            )  0           NUMBER OF SEMESTERS FOREIGN LANGUAGE MISSING
002   #FLANG-Y   (1-9          )  1           NUMBER OF SEMESTERS FOREIGN LANGUAGE NOT-MISSING

CONDITIONING ID:            BACK0037
DESCRIPTION:                NUMBER OF SEMESTERS FOREIGN LANGUAGE(S) (LINEAR)
GRADES/ASSESSMENTS:         N12
GROUP LABEL:                #LAN-LIN          LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                    B007105           DEGREES OF FREEDOM PER CONTRAST:  8
TYPE OF CONTRAST:           SCALE             NUMBER OF SPECIFICATION RECORDS: 10
001   #LAN-L01   (1            )  1           NUMBER OF SEMESTERS FOREIGN LANGUAGE LINEAR
```

Table F-2 (continued)
Conditioning Variables for Reading Main Samples

```
002   #LAN-L02   (2            )  2           NUMBER OF SEMESTERS FOREIGN LANGUAGE LINEAR
003   #LAN-L03   (3            )  3           NUMBER OF SEMESTERS FOREIGN LANGUAGE LINEAR
004   #LAN-L04   (4            )  4           NUMBER OF SEMESTERS FOREIGN LANGUAGE LINEAR
005   #LAN-L05   (5            )  5           NUMBER OF SEMESTERS FOREIGN LANGUAGE LINEAR
006   #LAN-L06   (6            )  6           NUMBER OF SEMESTERS FOREIGN LANGUAGE LINEAR
007   #LAN-L07   (7            )  7           NUMBER OF SEMESTERS FOREIGN LANGUAGE LINEAR
008   #LAN-L08   (8            )  8           NUMBER OF SEMESTERS FOREIGN LANGUAGE LINEAR
009   #LAN-L09   (9            )  9           NUMBER OF SEMESTERS FOREIGN LANGUAGE LINEAR
010   #LAN-L10   (M            )  0           NUMBER OF SEMESTERS FOREIGN LANGUAGE LINEAR (MISSING)

CONDITIONING ID:         BACK0038
DESCRIPTION:             NUMBER OF SEMESTERS VOC/TECH/BUSINESS EDUCATION (MISSING/NOT-MISSING)
GRADES/ASSESSMENTS:      N12
GROUP LABEL:             #SEM_VOC           LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                 B007106            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:        CLASS              NUMBER OF SPECIFICATION RECORDS:  2
001   #VOC-?     (M            )  0           NUMBER OF SEMESTERS VOC/TECH MISSING
002   #VOC-Y     (1-9          )  1           NUMBER OF SEMESTERS VOC/TECH NOT-MISSING

CONDITIONING ID:         BACK0039
DESCRIPTION:             NUMBER OF SEMESTERS VOCATIONAL/TECHNICAL/BUSINESS EDUCATION (LINEAR)
GRADES/ASSESSMENTS:      N12
GROUP LABEL:             #VOC-LIN           LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                 B007106            DEGREES OF FREEDOM PER CONTRAST:  8
TYPE OF CONTRAST:        SCALE              NUMBER OF SPECIFICATION RECORDS: 10
001   #VOC-L01   (1            )  1           NUMBER OF SEMESTERS VOC/TECH LINEAR
002   #VOC-L02   (2            )  2           NUMBER OF SEMESTERS VOC/TECH LINEAR
003   #VOC-L03   (3            )  3           NUMBER OF SEMESTERS VOC/TECH LINEAR
004   #VOC-L04   (4            )  4           NUMBER OF SEMESTERS VOC/TECH LINEAR
005   #VOC-L05   (5            )  5           NUMBER OF SEMESTERS VOC/TECH LINEAR
006   #VOC-L06   (6            )  6           NUMBER OF SEMESTERS VOC/TECH LINEAR
007   #VOC-L07   (7            )  7           NUMBER OF SEMESTERS VOC/TECH LINEAR
008   #VOC-L08   (8            )  8           NUMBER OF SEMESTERS VOC/TECH LINEAR
009   #VOC-L09   (9            )  9           NUMBER OF SEMESTERS VOC/TECH LINEAR
010   #VOC-L10   (M            )  0           NUMBER OF SEMESTERS VOC/TECH LINEAR (MISSING)

CONDITIONING ID:         BACK0040
DESCRIPTION:             NUMBER OF SEMESTERS ART/MUSIC (MISSING/NOT-MISSING)
GRADES/ASSESSMENTS:      N12
GROUP LABEL:             #SEM_ART           LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                 B007107            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:        CLASS              NUMBER OF SPECIFICATION RECORDS:  2
001   #ART-?     (M            )  0           NUMBER OF SEMESTERS ART/MUSIC MISSING
002   #ART-Y     (1-9          )  1           NUMBER OF SEMESTERS ART/MUSIC NOT-MISSING

CONDITIONING ID:         BACK0041
DESCRIPTION:             NUMBER OF SEMESTERS ART/MUSIC (LINEAR)
GRADES/ASSESSMENTS:      N12
GROUP LABEL:             #ART-LIN           LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                 B007107            DEGREES OF FREEDOM PER CONTRAST:  8
TYPE OF CONTRAST:        SCALE              NUMBER OF SPECIFICATION RECORDS: 10
001   #ART-L01   (1            )  1           NUMBER OF SEMESTERS ART/MUSIC LINEAR
002   #ART-L02   (2            )  2           NUMBER OF SEMESTERS ART/MUSIC LINEAR
003   #ART-L03   (3            )  3           NUMBER OF SEMESTERS ART/MUSIC LINEAR
004   #ART-L04   (4            )  4           NUMBER OF SEMESTERS ART/MUSIC LINEAR
005   #ART-L05   (5            )  5           NUMBER OF SEMESTERS ART/MUSIC LINEAR
006   #ART-L06   (6            )  6           NUMBER OF SEMESTERS ART/MUSIC LINEAR
007   #ART-L07   (7            )  7           NUMBER OF SEMESTERS ART/MUSIC LINEAR
008   #ART-L08   (8            )  8           NUMBER OF SEMESTERS ART/MUSIC LINEAR
009   #ART-L09   (9            )  9           NUMBER OF SEMESTERS ART/MUSIC LINEAR
010   #ART-L10   (M            )  0           NUMBER OF SEMESTERS ART/MUSIC LINEAR (MISSING)

CONDITIONING ID:         BACK0042
DESCRIPTION:             BORN IN ONE OF THE 50 STATES
GRADES/ASSESSMENTS:      N04, S04, N08, S08, N12
```

Table F-2 (continued)
Conditioning Variables for Reading Main Samples

```
GROUP LABEL:               BORN_USA          LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                   B007801           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:   2
001   USA-YES    (1           ) 0            BORN IN THE USA:  YES
002   USA-NO/?   (2,M         ) 1            BORN IN THE USA:  NO/MIS SING

CONDITIONING ID:           BACK0043
DESCRIPTION:               HOW MANY TIMES CHANGED SCHOOLS IN THE LAST TWO YEARS?
GRADES/ASSESSMENTS:        N04, S04, N08, S08
GROUP LABEL:               SCH_CHGS          LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                   B007301           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:   4
001   CHGSCH=0   (1           )  000         CHANGED SCHOOLS (NONE)
002   CHGSCH=1   (2           )  100         CHANGED SCHOOLS ONCE
003   CHGSCH=2   (3           )  010         CHANGED SCHOOLS TWICE
004   CHGSCH3+   (4,M         )  001         CHANGED SCHOOLS 3 OR MORE TIMES, MISSING

CONDITIONING ID:           BACK0044
DESCRIPTION:               HOW MANY GRADES HAVE YOU GONE TO SCHOOL IN THIS STATE? (K-4)
GRADES/ASSESSMENTS:        N04, S04
GROUP LABEL:               GRDS_ST4          LENGTH OF CONTRAST FIELD       :   2
NAEP ID:                   B007601           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:   3
001   ST4GRD<1   (1,M         )  00          LESS THAN ONE GRADE IN THIS STATE, MISSING (K-4)
002   ST4GRD12   (2           )  10          ONE TO TWO GRADES IN THIS STATE (K-4)
003   ST4GRD3+   (3           )  01          THREE OR MORE GRADES IN THIS STATE (K-4)

CONDITIONING ID:           BACK0045
DESCRIPTION:               HOW OFTEN DO YOU DISCUSS THINGS STUDIED IN SCHOOL WITH SOMEONE AT HOME?
GRADES/ASSESSMENTS:        N04, S04, N08, S08, N12
GROUP LABEL:               DISC@HOM          LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                   B007401           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:   4
001   DIS@HOM1   (1           )  000         DISCUSS AT HOME (ALMOST EVERYDAY)
002   DIS@HOM2   (2           )  100         DISCUSS AT HOME (ONCE OR TWICE A WEEK)
003   DIS@HOM3   (3           )  010         DISCUSS AT HOME (ONCE OR TWICE A MONTH)
004   DIS@HOM4   (4,M         )  001         DISCUSS AT HOME (NEVER OR HARDLY EVER, MISSING)

CONDITIONING ID:           BACK0046
DESCRIPTION:               HOW OFTEN DO USE A COMPUTER FOR SCHOOLWORK?
GRADES/ASSESSMENTS:        N04, S04, N08, S08, N12
GROUP LABEL:               COMP4SCH          LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                   B007501           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:   5
001   CMP4SCH1   (1           )  0000        COMPUTER FOR SCHOOLWORK (ALMOST EVERYDAY)
002   CMP4SCH2   (2           )  1000        COMPUTER FOR SCHOOLWORK (ONCE OR TWICE A WEEK)
003   CMP4SCH3   (3           )  0100        COMPUTER FOR SCHOOLWORK (ONCE OR TWICE A MONTH)
004   CMP4SCH4   (4           )  0010        COMPUTER FOR SCHOOLWORK (NEVER OR HARDLY EVER)
005   CMP4SCH5   (M           )  0001        COMPUTER FOR SCHOOLWORK (MISSING)

CONDITIONING ID:           BACK0047
DESCRIPTION:               HOW MANY GRADES HAVE YOU GONE TO SCHOOL IN THIS STATE? (K-8)
GRADES/ASSESSMENTS:        N08, S08
GROUP LABEL:               GRDS_ST8          LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                   B007701           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:   4
001   ST8GRD<1   (1,M         )  000         LESS THAN ONE GRADE IN THIS STATE, MISSING (K-8)
002   ST8GRD12   (2           )  100         ONE TO TWO GRADES IN THIS STATE (K-8)
003   ST8GRD35   (3           )  010         THREE TO FIVE GRADES IN THIS STATE (K-8)
004   ST8GRD>5   (4           )  001         MORE THAN FIVE GRADES IN THIS STATE (K-8)

CONDITIONING ID:           BACK0048
DESCRIPTION:               WHAT DO YOU EXPECT YOUR MAIN ACTIVITY WILL BE THE YEAR AFTER SCHOOL?
GRADES/ASSESSMENTS:        N12
GROUP LABEL:               MAINACT           LENGTH OF CONTRAST FIELD       :   5
```

BEST COPY AVAILABLE

Table F-2 (continued)
Conditioning Variables for Reading Main Samples

```
NAEP ID:                    B007201          DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS            NUMBER OF SPECIFICATION RECORDS:   6
001   WORK_F/T    (1          )  00000       MAIN ACTIVITY:  WORK FULL-TIME
002   VOC/BUSI    (2          )  10000       MAIN ACTIVITY:  ATTEND VOC, TECH, BUSINESS SCHOOL
003   2-YR_COL    (3          )  01000       MAIN ACTIVITY:  2-YEAR COLLEGE
004   4-YR_COL    (4          )  00100       MAIN ACTIVITY:  4-YEAR COLLEGE
005   MILITARY    (5          )  00010       MAIN ACTIVITY:  MILITARY
006   OTHERACT    (6,M        )  00001       MAIN ACTIVITY:  OTHER, MISSING

CONDITIONING ID:            READ0001
DESCRIPTION:                SCHOOL LEVEL AVERAGE READING PROFICIENCY
GRADES/ASSESSMENTS:         N04, S04, N08, S08, N12
GROUP LABEL:                SLP_READ         LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                    SCHREAD          DEGREES OF FREEDOM PER CONTRAST: 999
TYPE OF CONTRAST:           CLASS            NUMBER OF SPECIFICATION RECORDS:   2
001   SLP_RD-Y    (@          )  1           SCHOOL LEVEL AVERAGE READING PROFICIENCY NOT-MISSING
002   SLP_RD-?    (M          )  0           SCHOOL LEVEL AVERAGE READING PROFICIENCY MISSING

CONDITIONING ID:            READ0002
DESCRIPTION:                SCHOOL LEVEL AVERAGE READING PROFICIENCY
GRADES/ASSESSMENTS:         N04, S04, N08, S08, N12
GROUP LABEL:                SLP_REDL         LENGTH OF CONTRAST FIELD       :   8
NAEP ID:                    SCHREAD          DEGREES OF FREEDOM PER CONTRAST: 999
TYPE OF CONTRAST:           SCALE            NUMBER OF SPECIFICATION RECORDS:   2
001   SLP_RD-L    (#          )  (F8.4)      SCHOOL LEVEL AVERAGE READING PROFICIENCY MEAN
002   SLP_RD-L    (M          )  0           SCHOOL LEVEL AVERAGE READING PROFICIENCY MISSING

CONDITIONING ID:            READ0003
DESCRIPTION:                DURING PAST MONTH, HOW MANY BOOKS HAVE YOU READ ON YOUR OWN OUTSIDE OF SCHOOL?
GRADES/ASSESSMENTS:         N04, S04, N08, N12
GROUP LABEL:                NBOOKSRD         LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                    R810801          DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS            NUMBER OF SPECIFICATION RECORDS:   5
001   NBOOKS-1    (1          )  0000        NUMBER OF BOOKS READ:  NONE
002   NBOOKS-2    (2          )  1000        NUMBER OF BOOKS READ:  ONE OR TWO
003   NBOOKS-3    (3          )  0100        NUMBER OF BOOKS READ:  THREE OR FOUR
004   NBOOKS-4    (4          )  0010        NUMBER OF BOOKS READ:  FIVE OR MORE
005   NBOOKS-?    (M          )  0001        NUMBER OF BOOKS READ:  MISSING

CONDITIONING ID:            READ0004
DESCRIPTION:                WHAT KIND OF READER DO YOU THINK YOU ARE?
GRADES/ASSESSMENTS:         N04, S04, N08, N12
GROUP LABEL:                KIND_RDR         LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                    R810201          DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS            NUMBER OF SPECIFICATION RECORDS:   5
001   READ-VGD    (1          )  0000        KIND OF READER:  A VERY GOOD READER
002   READ-GD     (2          )  1000        KIND OF READER:  A GOOD READER
003   READ-AVG    (3          )  0100        KIND OF READER:  AN AVERAGE READER
004   READ-PR     (4          )  0010        KIND OF READER:  A POOR READER
005   READ-?      (M          )  0001        KIND OF READER:  MISSING

CONDITIONING ID:            READ0005
DESCRIPTION:                HOW OFTEN DO YOU READ FOR FUN ON YOUR OWN TIME?
GRADES/ASSESSMENTS:         N04, S04, N08, N12
GROUP LABEL:                READ4FUN         LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                    R810901          DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS            NUMBER OF SPECIFICATION RECORDS:   5
001   RD4FUN-1    (1          )  0000        READ FOR FUN:  ALMOST EVERY DAY
002   RD4FUN-2    (2          )  1000        READ FOR FUN:  ONCE OR TWICE A WEEK
003   RD4FUN-3    (3          )  0100        READ FOR FUN:  ONCE OR TWICE A MONTH
004   RD4FUN-4    (4          )  0010        READ FOR FUN:  NEVER OR HARDLY EVER
005   RD4FUN-?    (M          )  0001        READ FOR FUN:  MISSING

CONDITIONING ID:            READ0006
DESCRIPTION:                HOW OFTEN DO YOU READ A STORY OR NOVEL?
```

Table F-2 (continued)
Conditioning Variables for Reading Main Samples

```
GRADES/ASSESSMENTS:        N08, N12
GROUP LABEL:               RD_NOVEL           LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                   R810904            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS              NUMBER OF SPECIFICATION RECORDS:  5
001   RDNOVL-1   (1           )  0000         READ STORY/NOVEL:  ALMOST EVERY DAY
002   RDNOVL-2   (2           )  1000         READ STORY/NOVEL:  ONCE OR TWICE A WEEK
003   RDNOVL-3   (3           )  0100         READ STORY/NOVEL:  ONCE OR TWICE A MONTH
004   RDNOVL-4   (4           )  0010         READ STORY/NOVEL:  NEVER OR HARDLY EVER
005   RDNOVL-?   (M           )  0001         READ STORY/NOVEL:  MISSING

CONDITIONING ID:           READ0007
DESCRIPTION:               HOW OFTEN DO YOU READ A NEWSPAPER?
GRADES/ASSESSMENTS:        N08, N12
GROUP LABEL:               RD_NEWSP           LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                   R810905            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS              NUMBER OF SPECIFICATION RECORDS:  5
001   RDNEWP-1   (1           )  0000         READ NEWSPAPER:  ALMOST EVERY DAY
002   RDNEWP-2   (2           )  1000         READ NEWSPAPER:  ONCE OR TWICE A WEEK
003   RDNEWP-3   (3           )  0100         READ NEWSPAPER:  ONCE OR TWICE A MONTH
004   RDNEWP-4   (4           )  0010         READ NEWSPAPER:  NEVER OR HARDLY EVER
005   RDNEWP-?   (M           )  0001         READ NEWSPAPER:  MISSING

CONDITIONING ID:           READ0008
DESCRIPTION:               HOW OFTEN DO YOU READ A MAGAZINE?
GRADES/ASSESSMENTS:        N08, N12
GROUP LABEL:               RD_MAGAZ           LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                   R810906            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS              NUMBER OF SPECIFICATION RECORDS:  5
001   RDMAGZ-1   (1           )  0000         READ MAGAZINE:  ALMOST EVERY DAY
002   RDMAGZ-2   (2           )  1000         READ MAGAZINE:  ONCE OR TWICE A WEEK
003   RDMAGZ-3   (3           )  0100         READ MAGAZINE:  ONCE OR TWICE A MONTH
004   RDMAGZ-4   (4           )  0010         READ MAGAZINE:  NEVER OR HARDLY EVER
005   RDMAGZ-?   (M           )  0001         READ MAGAZINE:  MISSING

CONDITIONING ID:           READ0009
DESCRIPTION:               HOW OFTEN DO YOU TALK WITH YOUR FRIENDS OR FAMILY ABOUT SOMTHING YOU HAVE READ?
GRADES/ASSESSMENTS:        N04, S04, N08, N12
GROUP LABEL:               TALKREAD           LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                   R810902            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS              NUMBER OF SPECIFICATION RECORDS:  5
001   TALKRD-1   (1           )  0000         TALK ABOUT READING: ALMOST EVERY DAY
002   TALKRD-2   (2           )  1000         TALK ABOUT READING: ONCE OR TWICE A WEEK
003   TALKRD-3   (3           )  0100         TALK ABOUT READING: ONCE OR TWICE A MONTH
004   TALKRD-4   (4           )  0010         TALK ABOUT READING: NEVER OR HARDLY EVER
005   TALKRD-?   (M           )  0001         TALK ABOUT READING: MISSING

CONDITIONING ID:           READ0010
DESCRIPTION:               HOW OFTEN DO YOU TAKE BOOKS OUT OF THE LIBRARY FOR YOU OWN ENJOYMENT?
GRADES/ASSESSMENTS:        N04, S04, N08, N12
GROUP LABEL:               USELIBRY           LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                   R810903            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS              NUMBER OF SPECIFICATION RECORDS:  5
001   USELIB-1   (1           )  0000         USE THE LIBRARY:  ALMOST EVERY DAY
002   USELIB-2   (2           )  1000         USE THE LIBRARY:  ONCE OR TWICE A WEEK
003   USELIB-3   (3           )  0100         USE THE LIBRARY:  ONCE OR TWICE A MONTH
004   USELIB-4   (4           )  0010         USE THE LIBRARY:  NEVER OR HARDLY EVER
005   USELIB-?   (M           )  0001         USE THE LIBRARY:  MISSING

CONDITIONING ID:           READ0011
DESCRIPTION:               HOW OFTEN DOES YOUR TEACHER DISCUSS NEW OR DIFFICULT VOCABULARY?
GRADES/ASSESSMENTS:        N04, S04, N08, N12
GROUP LABEL:               S_VOCAB            LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                   R811001            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS              NUMBER OF SPECIFICATION RECORDS:  5
001   VOCAB-S1   (1           )  0000         DISCUSS VOCABULARY:  ALMOST EVERY DAY
```

Table F-2 (continued)
Conditioning Variables for Reading Main Samples

```
002   VOCAB-S2   (2           )  1000        DISCUSS VOCABULARY:  ONCE OR TWICE A WEEK
003   VOCAB-S3   (3           )  0100        DISCUSS VOCABULARY:  ONCE OR TWICE A MONTH
004   VOCAB-S4   (4           )  0010        DISCUSS VOCABULARY:  NEVER OR HARDLY EVER
005   VOCAB-S?   (M           )  0001        DISCUSS VOCABULARY:  MISSING

CONDITIONING ID:             READ0012
DESCRIPTION:                 HOW OFTEN YOUR TEACHER ASKS STUDENTS TO TALK TO EACH OTHER ABOUT WHAT READ?
GRADES/ASSESSMENTS:          N04, S04, N08, N12
GROUP LABEL:                 S_TALKRD            LENGTH OF CONTRAST FIELD        :   4
NAEP ID:                     R811002             DEGREES OF FREEDOM PER CONTRAST:    1
TYPE OF CONTRAST:            CLASS               NUMBER OF SPECIFICATION RECORDS:    5
001   TLKRD-S1   (1           )  0000        TEACHER ASK TO TALK ABOUT READING:  ALMOST EVERY DAY
002   TLKRD-S2   (2           )  1000        TEACHER ASK TO TALK ABOUT READING:  ONCE OR TWICE A WEEK
003   TLKRD-S3   (3           )  0100        TEACHER ASK TO TALK ABOUT READING:  ONCE OR TWICE A MONTH
004   TLKRD-S4   (4           )  0010        TEACHER ASK TO TALK ABOUT READING:  NEVER OR HARDLY EVER
005   TLKRD-S?   (M           )  0001        TEACHER ASK TO TALK ABOUT READING:  MISSING

CONDITIONING ID:             READ0013
DESCRIPTION:                 HOW OFTEN DOES TEACHER ASK YOU TO WORK IN A READING WORKBOOK OR ON A WORKSHEET?
GRADES/ASSESSMENTS:          N04, S04, N08, N12
GROUP LABEL:                 S_WBKWSH            LENGTH OF CONTRAST FIELD        :   4
NAEP ID:                     R811003             DEGREES OF FREEDOM PER CONTRAST:    1
TYPE OF CONTRAST:            CLASS               NUMBER OF SPECIFICATION RECORDS:    5
001   WB/WS-S1   (1           )  0000        READING WORKBOOK/WORKSHEET:  ALMOST EVERY DAY
002   WB/WS-S2   (2           )  1000        READING WORKBOOK/WORKSHEET:  ONCE OR TWICE A WEEK
003   WB/WS-S3   (3           )  0100        READING WORKBOOK/WORKSHEET:  ONCE OR TWICE A MONTH
004   WB/WS-S4   (4           )  0010        READING WORKBOOK/WORKSHEET:  NEVER OR HARDLY EVER
005   WB/WS-S?   (M           )  0001        READING WORKBOOK/WORKSHEET:  MISSING

CONDITIONING ID:             READ0014
DESCRIPTION:                 HOW OFTEN DOES YOUR TEACHER ASK YOU TO WRITE SOMETHING ABOUT WHAT YOU HAVE READ?
GRADES/ASSESSMENTS:          N04, S04, N08, N12
GROUP LABEL:                 S_WRITRD            LENGTH OF CONTRAST FIELD        :   4
NAEP ID:                     R811004             DEGREES OF FREEDOM PER CONTRAST:    1
TYPE OF CONTRAST:            CLASS               NUMBER OF SPECIFICATION RECORDS:    5
001   WRTRD-S1   (1           )  0000        WRITE ABOUT READING:  ALMOST EVERY DAY
002   WRTRD-S2   (2           )  1000        WRITE ABOUT READING:  ONCE OR TWICE A WEEK
003   WRTRD-S3   (3           )  0100        WRITE ABOUT READING:  ONCE OR TWICE A MONTH
004   WRTRD-S4   (4           )  0010        WRITE ABOUT READING:  NEVER OR HARDLY EVER
005   WRTRD-S?   (M           )  0001        WRITE ABOUT READING:  MISSING

CONDITIONING ID:             READ0015
DESCRIPTION:                 HOW OFTEN DOES TCHER ASK STUDENTS TO DO GROUP ACTIVITY/PROJECT ABOUT WHAT READ?
GRADES/ASSESSMENTS:          N04, S04, N08, N12
GROUP LABEL:                 S_RDPROJ            LENGTH OF CONTRAST FIELD        :   4
NAEP ID:                     R811005             DEGREES OF FREEDOM PER CONTRAST:    1
TYPE OF CONTRAST:            CLASS               NUMBER OF SPECIFICATION RECORDS:    5
001   RDPRJ-S1   (1           )  0000        PROJECT ABOUT READING: ALMOST EVERY DAY
002   RDPRJ-S2   (2           )  1000        PROJECT ABOUT READING: ONCE OR TWICE A WEEK
003   RDPRJ-S3   (3           )  0100        PROJECT ABOUT READING: ONCE OR TWICE A MONTH
004   RDPRJ-S4   (4           )  0010        PROJECT ABOUT READING: NEVER OR HARDLY EVER
005   RDPRJ-S?   (M           )  0001        PROJECT ABOUT READING: MISSING

CONDITIONING ID:             READ0016
DESCRIPTION:                 HOW OFTEN DOES YOUR TEACHER ASK STUDENTS TO READ ALOUD?
GRADES/ASSESSMENTS:          N04, S04, N08, N12
GROUP LABEL:                 S_ALOUD             LENGTH OF CONTRAST FIELD        :   4
NAEP ID:                     R811006             DEGREES OF FREEDOM PER CONTRAST:    1
TYPE OF CONTRAST:            CLASS               NUMBER OF SPECIFICATION RECORDS:    5
001   ALOUD-S1   (1           )  0000        READ ALOUD: ALMOST EVERY DAY
002   ALOUD-S2   (2           )  1000        READ ALOUD: ONCE OR TWICE A WEEK
003   ALOUD-S3   (3           )  0100        READ ALOUD: ONCE OR TWICE A MONTH
004   ALOUD-S4   (4           )  0010        READ ALOUD: NEVER OR HARDLY EVER
005   ALOUD-S?   (M           )  0001        READ ALOUD: MISSING
```

Table F-2 (continued)
Conditioning Variables for Reading Main Samples

```
CONDITIONING ID:              READ0017
DESCRIPTION:                  HOW OFTEN DOES YOUR TEACHER ASK YOU TO READ SILENTLY?
GRADES/ASSESSMENTS:           N04, S04, N08, N12
GROUP LABEL:                  S_SILENT          LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                      R811007           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:             CLASS             NUMBER OF SPECIFICATION RECORDS:  5
001    SILNT-S1    (1            ) 0000         READ SILENTLY: ALMOST EVERY DAY
002    SILNT-S2    (2            ) 1000         READ SILENTLY: ONCE OR TWICE A WEEK
003    SILNT-S3    (3            ) 0100         READ SILENTLY: ONCE OR TWICE A MONTH
004    SILNT-S4    (4            ) 0010         READ SILENTLY: NEVER OR HARDLY EVER
005    SILNT-S?    (M            ) 0001         READ SILENTLY: MISSING

CONDITIONING ID:              READ0018
DESCRIPTION:                  HOW OFTEN DOES TEACHER ASK YOU TO WRITE IN A LOG OR JOURNAL ABOUT WHAT YOU READ?
GRADES/ASSESSMENTS:           N04, S04, N08, N12
GROUP LABEL:                  S_RDLOG           LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                      R811008           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:             CLASS             NUMBER OF SPECIFICATION RECORDS:  5
001    RDLOG-S1    (1            ) 0000         WRITE IN LOG/JOURNAL:  ALMOST EVERY DAY
002    RDLOG-S2    (2            ) 1000         WRITE IN LOG/JOURNAL:  ONCE OR TWICE A WEEK
003    RDLOG-S3    (3            ) 0100         WRITE IN LOG/JOURNAL:  ONCE OR TWICE A MONTH
004    RDLOG-S4    (4            ) 0010         WRITE IN LOG/JOURNAL:  NEVER OR HARDLY EVER
005    RDLOG-S?    (M            ) 0001         WRITE IN LOG/JOURNAL:  MISSING

CONDITIONING ID:              READ0019
DESCRIPTION:                  HOW OFTEN DOES TEACHER GIVE YOU TIME TO READ BOOKS YOU HAVE CHOSEN YOURSELF?
GRADES/ASSESSMENTS:           N04, S04, N08, N12
GROUP LABEL:                  S_OWNBKS          LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                      R811009           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:             CLASS             NUMBER OF SPECIFICATION RECORDS:  5
001    OWNBK-S1    (1            ) 0000         BOOKS CHOSEN YOURSELF:  ALMOST EVERY DAY
002    OWNBK-S2    (2            ) 1000         BOOKS CHOSEN YOURSELF:  OR TWICE A WEEK
003    OWNBK-S3    (3            ) 0100         BOOKS CHOSEN YOURSELF:  ONCE OR TWICE A MONTH
004    OWNBK-S4    (4            ) 0010         BOOKS CHOSEN YOURSELF:  NEVER OR HARDLY EVER
005    OWNBK-S?    (M            ) 0001         BOOKS CHOSEN YOURSELF:  MISSING

CONDITIONING ID:              READ0020
DESCRIPTION:                  HOW OFTEN DOES TEACHER ASK YOU TO EXPLAIN OR SUPPORT UNDERSTANDING OF WHAT READ?
GRADES/ASSESSMENTS:           N08, N12
GROUP LABEL:                  EXPLAIN           LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                      R811010           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:             CLASS             NUMBER OF SPECIFICATION RECORDS:  5
001    EXPLA-1     (1            ) 0000         EXPLAIN WHAT YOU READ:  ALMOST EVERY DAY
002    EXPLA-2     (2            ) 1000         EXPLAIN WHAT YOU READ:  OR TWICE A WEEK
003    EXPLA-3     (3            ) 0100         EXPLAIN WHAT YOU READ:  ONCE OR TWICE A MONTH
004    EXPLA-4     (4            ) 0010         EXPLAIN WHAT YOU READ:  NEVER OR HARDLY EVER
005    EXPLA-?     (M            ) 0001         EXPLAIN WHAT YOU READ:  MISSING

CONDITIONING ID:              READ0021
DESCRIPTION:                  HOW OFTEN DOES TEACHER ASK YOU TO DISCUSS INTERPRETATIONS OF WHAT YOU HAVE READ?
GRADES/ASSESSMENTS:           N08, N12
GROUP LABEL:                  DISCUSS           LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                      R811011           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:             CLASS             NUMBER OF SPECIFICATION RECORDS:  5
001    DISCU-1     (1            ) 0000         DISCUSS INTERPRETATIONS OF READING:  ALMOST EVERY DAY
002    DISCU-2     (2            ) 1000         DISCUSS INTERPRETATIONS OF READING:  OR TWICE A WEEK
003    DISCU-3     (3            ) 0100         DISCUSS INTERPRETATIONS OF READING:  ONCE OR TWICE A MONTH
004    DISCU-4     (4            ) 0010         DISCUSS INTERPRETATIONS OF READING:  NEVER OR HARDLY EVER
005    DISCU-?     (M            ) 0001         DISCUSS INTERPRETATIONS OF READING:  MISSING

CONDITIONING ID:              READ0022
DESCRIPTION:                  HOW OFTEN DOES YOUR TEACHER ASK YOU TO MAKE PREDICTIONS AS YOU CONTINUE READING?
GRADES/ASSESSMENTS:           N08, N12
GROUP LABEL:                  PREDICT           LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                      R811012           DEGREES OF FREEDOM PER CONTRAST:  1
```

Table F-2 (continued)
Conditioning Variables for Reading Main Samples

```
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:   5
001   PREDI-1    (1          )  0000          PREDICTIONS OF READING:  ALMOST EVERY DAY
002   PREDI-2    (2          )  1000          PREDICTIONS OF READING:  OR TWICE A WEEK
003   PREDI-3    (3          )  0100          PREDICTIONS OF READING:  ONCE OR TWICE A MONTH
004   PREDI-4    (4          )  0010          PREDICTIONS OF READING:  NEVER OR HARDLY EVER
005   PREDI-?    (M          )  0001          PREDICTIONS OF READING:  MISSING

CONDITIONING ID:           READ0023
DESCRIPTION:               HOW OFTEN DO YOU USE A LIBRARY TO DO RESEARCH FOR A SCHOOL ASSIGNMENT?
GRADES/ASSESSMENTS:        N08, N12
GROUP LABEL:               LIB4RSCH          LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                   R811301           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:   5
001   LIBRS-1    (1          )  0000          LIBRARY FOR RESEARCH:  ALMOST EVERY DAY
002   LIBRS-2    (2          )  1000          LIBRARY FOR RESEARCH:  ONCE OR TWICE A WEEK
003   LIBRS-3    (3          )  0100          LIBRARY FOR RESEARCH:  ONCE OR TWICE A MONTH
004   LIBRS-4    (4,5        )  0010          LIBRARY FOR RESEARCH:  ONCE OR TWICE/YEAR, NEVER/HARDLY EVER
005   LIBRS-?    (M          )  0001          LIBRARY FOR RESEARCH:  MISSING

CONDITIONING ID:           READ0024
DESCRIPTION:               HOW OFTEN DO YOU USE A LIBRARY TO BORROW BOOKS FOR A SCHOOL ASSIGNMENT?
GRADES/ASSESSMENTS:        N08, N12
GROUP LABEL:               LIB2BORW          LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                   R811302           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:   5
001   LIBBR-1    (1          )  0000          LIBRARY TO BORROW BOOKS:  ALMOST EVERY DAY
002   LIBBR-2    (2          )  1000          LIBRARY TO BORROW BOOKS:  OR TWICE A WEEK
003   LIBBR-3    (3          )  0100          LIBRARY TO BORROW BOOKS:  ONCE OR TWICE A MONTH
004   LIBBR-4    (4,5        )  0010          LIBRARY TO BORROW BKS:  ONCE OR TWICE/YEAR, NEVER/HARD EVER
005   LIBBR-?    (M          )  0001          LIBRARY TO BORROW BOOKS:  MISSING

CONDITIONING ID:           READ0025
DESCRIPTION:               HOW OFTEN DO YOU USE A LIBRARY FOR INFORMATION FOR YOUR OWN USE?
GRADES/ASSESSMENTS:        N08, N12
GROUP LABEL:               LIB4INFO          LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                   R811303           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:   5
001   LIBIN-1    (1          )  0000          LIBRARY FOR INFORMATION:  ALMOST EVERY DAY
002   LIBIN-2    (2          )  1000          LIBRARY FOR INFORMATION:  OR TWICE A WEEK
003   LIBIN-3    (3          )  0100          LIBRARY FOR INFORMATION:  ONCE OR TWICE A MONTH
004   LIBIN-4    (4,5        )  0010          LIBRARY FOR INFO:  ONCE OR TWICE A YEAR, NEVER/HARDLY EVER
005   LIBIN-?    (M          )  0001          LIBRARY FOR INFORMATION:  MISSING

CONDITIONING ID:           READ0026
DESCRIPTION:               HOW OFTEN DO YOU USE A LIBRARY AS A QUIET PLACE TO STUDY?
GRADES/ASSESSMENTS:        N08, N12
GROUP LABEL:               LIBQUIET          LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                   R811304           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:   5
001   LIBQU-1    (1          )  0000          LIBRARY FOR QUIET STUDY:  ALMOST EVERY DAY
002   LIBQU-2    (2          )  1000          LIBRARY FOR QUIET STUDY:  OR TWICE A WEEK
003   LIBQU-3    (3          )  0100          LIBRARY FOR QUIET STUDY:  ONCE OR TWICE A MONTH
004   LIBQU-4    (4,5        )  0010          LIB FOR QUIET STUDY:  ONCE OR TWICE/YEAR, NEVER/HARDLY EVER
005   LIBQU-?    (M          )  0001          LIBRARY FOR QUIET STUDY:  MISSING

CONDITIONING ID:           READ0027
DESCRIPTION:               ABOUT HOW MANY QUESTIONS DID YOU GET RIGHT ON THE READING TEST?
GRADES/ASSESSMENTS:        N04, S04, N08, N12
GROUP LABEL:               #QUESTN+          LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                   RM00101           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:   4
001   #QUEST+1   (1          )  000           NUMBER QUESTIONS RIGHT:  ALMOST ALL
002   #QUEST+2   (2          )  100           NUMBER QUESTIONS RIGHT:  MORE THAN HALF
003   #QUEST+3   (3          )  010           NUMBER QUESTIONS RIGHT:  ABOUT HALF
004   #QUEST+4   (4,M        )  001           NUMBER QUESTIONS RIGHT:  LESS THAN HALF, MISSING
```

Table F-2 (continued)
Conditioning Variables for Reading Main Samples

```
CONDITIONING ID:           READ0028
DESCRIPTION:               HOW HARD WAS THIS READING TEST COMPARED TO OTHERS?
GRADES/ASSESSMENTS:        N04, S04, N08, N12
GROUP LABEL:               TEST_DIF          LENGTH OF CONTRAST FIELD      :  4
NAEP ID:                   RM00201           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:  5
001   TESTDIF1   (1          )  0000         TEST DIFFICULTY:  MUCH HARDER THAN OTHERS
002   TESTDIF2   (2          )  1000         TEST DIFFICULTY:  HARDER THAN OTHERS
003   TESTDIF3   (3          )  0100         TEST DIFFICULTY:  ABOUT AS HARD AS OTHERS
004   TESTDIF4   (4          )  0010         TEST DIFFICULTY:  EASIER THAN OTHERS
005   TESTDIF?   (M          )  0001         TEST DIFFICULTY:  MISSING

CONDITIONING ID:           READ0029
DESCRIPTION:               HOW HARD DID YOU TRY ON THIS TEST COMPARED TO OTHER READING TESTS?
GRADES/ASSESSMENTS:        N04, S04, N08, N12
GROUP LABEL:               TEST_EFF          LENGTH OF CONTRAST FIELD      :  4
NAEP ID:                   RM00301           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:  5
001   TESTEFF1   (1          )  0000         TEST EFFORT:  MUCH HARDER THAN OTHERS
002   TESTEFF2   (2          )  1000         TEST EFFORT:  HARDER THAN OTHERS
003   TESTEFF3   (3          )  0100         TEST EFFORT:  ABOUT AS HARD AS OTHERS
004   TESTEFF4   (4          )  0010         TEST EFFORT:  NOT AS HARD AS OTHERS
005   TESTEFF?   (M          )  0001         TEST EFFORT:  MISSING

CONDITIONING ID:           READ0030
DESCRIPTION:               HOW IMPORTANT WAS IT TO YOU TO DO WELL ON THE READING TEST?
GRADES/ASSESSMENTS:        N04, S04, N08, N12
GROUP LABEL:               TEST_IMP          LENGTH OF CONTRAST FIELD      :  4
NAEP ID:                   RM00401           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:  5
001   TESTIMP1   (1          )  0000         TEST IMPORTANCE:  VERY IMPORTANT
002   TESTIMP2   (2          )  1000         TEST IMPORTANCE:  IMPORTANT
003   TESTIMP3   (3          )  0100         TEST IMPORTANCE:  SOMEWHAT IMPORTANT
004   TESTIMP4   (4          )  0010         TEST IMPORTANCE:  NOT VERY IMPORTANT
005   TESTIMP?   (M          )  0001         TEST IMPORTANCE:  MISSING

CONDITIONING ID:           READ0031
DESCRIPTION:               HOW OFTEN WERE YOU ASKED TO WRITE LONG ANSWERS ON READING TESTS?
GRADES/ASSESSMENTS:        N04, S04, N08, N12
GROUP LABEL:               LONG_ANS          LENGTH OF CONTRAST FIELD      :  4
NAEP ID:                   RM00501           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:  5
001   DSOLUTN1   (1          )  0000         DETAILED SOLUTIONS:  AT LEAST ONCE A WEEK
002   DSOLUTN2   (2          )  1000         DETAILED SOLUTIONS:  ONCE OR TWICE A MONTH
003   DSOLUTN3   (3          )  0100         DETAILED SOLUTIONS:  ONCE OR TWICE A YEAR
004   DSOLUTN4   (4          )  0010         DETAILED SOLUTIONS:  NEVER
005   DSOLUTN5   (M          )  0001         DETAILED SOLUTIONS:  MISSING

CONDITIONING ID:           SCHL0001
DESCRIPTION:               WHO TEACHES ENGLISH/LANGUAGE ARTS TO 8TH GRADE?
GRADES/ASSESSMENTS:        N08, S08
GROUP LABEL:               TSUB_ENG          LENGTH OF CONTRAST FIELD      :  2
NAEP ID:                   C034701           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:  3
001   ENG-T>1S   (1,M        )  00           8TH-GRADE ENGLISH:  TEACHERS WITH MORE THAN ONE SUBJECT
002   ENG-T=1S   (2          )  10           8TH-GRADE ENGLISH:  TEACHERS WITH ONE SUBJECT
003   ENG-SNT    (3          )  01           8TH-GRADE ENGLISH:  SUBJECT NOT TAUGHT

CONDITIONING ID:           SCHL0002
DESCRIPTION:               HAS READING BEEN IDENTIFIED AS A PRIORITY? (GRADE 4)
GRADES/ASSESSMENTS:        N04, S04, N08, S08
GROUP LABEL:               PRIOR-RD          LENGTH OF CONTRAST FIELD      :  2
NAEP ID:                   C031601           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:  3
```

Table F-2 (continued)
Conditioning Variables for Reading Main Samples

```
001   RPRIOR-Y   (1           )  00         READING PRIORITY:  YES
002   RPRIOR-N   (2           )  10         READING PRIORITY:  NO
003   RPRIOR-?   (M           )  01         READING PRIORITY:  MISSING

CONDITIONING ID:             SCHL0003
DESCRIPTION:                 HAS WRITING BEEN IDENTIFIED AS A PRIORITY? (GRADE 4)
GRADES/ASSESSMENTS:          N04, S04, N08, S08
GROUP LABEL:                 PRIOR-WR           LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                     C031602            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS              NUMBER OF SPECIFICATION RECORDS:  3
001   WPRIOR-Y   (1           )  00         WRITING PRIORITY:  YES
002   WPRIOR-N   (2           )  10         WRITING PRIORITY:  NO
003   WPRIOR-?   (M           )  01         WRITING PRIORITY:  MISSING

CONDITIONING ID:             SCHL0004
DESCRIPTION:                 WHAT PERCENT OF STUDENTS RECEIVE SUBSIDIZED LUNCH?
GRADES/ASSESSMENTS:          N04, S04, N08, S08, N12
GROUP LABEL:                 %SUBLUN            LENGTH OF CONTRAST FIELD       :  5
NAEP ID:                     C032001            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS              NUMBER OF SPECIFICATION RECORDS:  6
001   %SUBLUN1   (1,2,3       )  00000      PERCENT SUBSIDIZED LUNCH:  NONE-10%
002   %SUBLUN2   (4           )  10000      PERCENT SUBSIDIZED LUNCH:  11-25%
003   %SUBLUN3   (5           )  01000      PERCENT SUBSIDIZED LUNCH:  26-50%
004   %SUBLUN4   (6           )  00100      PERCENT SUBSIDIZED LUNCH:  51-75%
005   %SUBLUN5   (7,8         )  00010      PERCENT SUBSIDIZED LUNCH:  76-100%
006   %SUBLUN?   (M           )  00001      PERCENT SUBSIDIZED LUNCH:  MISSING

CONDITIONING ID:             SCHL0005
DESCRIPTION:                 WHAT PERCENT OF STUDENTS RECEIVE REMEDIAL READING?
GRADES/ASSESSMENTS:          N04, S04, N08, S08, N12
GROUP LABEL:                 %REMDL-R           LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                     C032002            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS              NUMBER OF SPECIFICATION RECORDS:  5
001   %REMRED1   (1,2         )  0000       PERCENT REMEDIAL READING:  NONE-5%
002   %REMRED2   (3           )  1000       PERCENT REMEDIAL READING:  6-10%
003   %REMRED3   (4           )  0100       PERCENT REMEDIAL READING:  11-25%
004   %REMRED4   (5,6,7,8     )  0010       PERCENT REMEDIAL READING:  26-100%
005   %REMRED?   (M           )  0001       PERCENT REMEDIAL READING:  MISSING

CONDITIONING ID:             SCHL0006
DESCRIPTION:                 WHAT PERCENTAGE OF STUDENTS ARE ENROLLED AT BEGINNING AND END OF SCHOOL YEAR?
GRADES/ASSESSMENTS:          N04, S04, N08, S08, N12
GROUP LABEL:                 %ENR/YR            LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                     C033700            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS              NUMBER OF SPECIFICATION RECORDS:  5
001   %ENR/YR1   (1           )  0000       YEAR LONG ENROLLMENT:  98-100 PERCENT
002   %ENR/YR2   (2           )  1000       YEAR LONG ENROLLMENT:  95-97 PERCENT
003   %ENR/YR3   (3           )  0100       YEAR LONG ENROLLMENT:  90-94 PERCENT
004   %ENR/YR4   (4           )  0010       YEAR LONG ENROLLMENT:  LESS THAN 90 PERCENT
005   %ENR/YR?   (M           )  0001       YEAR LONG ENROLLMENT:  MISSING

CONDITIONING ID:             SCHL0007
DESCRIPTION:                 HOW IS 4TH GRADE ORGANIZED AT YOUR SCHOOL?
GRADES/ASSESSMENTS:          N04, S04
GROUP LABEL:                 ORGANIZ4           LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                     C030900            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS              NUMBER OF SPECIFICATION RECORDS:  4
001   SELFCONT   (1           )  000        4TH GRADE ORGANIZATION:  SELF CONTAINED
002   DEPTLIZD   (2           )  100        4TH GRADE ORGANIZATION:  DEPARTMENTALIZED
003   REGRPED    (3           )  010        4TH GRADE ORGANIZATION:  REGROUPED
004   ORGANIZ?   (M           )  001        4TH GRADE ORGANIZATION:  MISSING

CONDITIONING ID:             SCHL0008
DESCRIPTION:                 HOW IS 8TH GRADE ORGANIZED AT YOUR SCHOOL?
GRADES/ASSESSMENTS:          N08, S08
```

Table F-2 (continued)
Conditioning Variables for Reading Main Samples

```
GROUP LABEL:                ORGANIZ8            LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                    C034200             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   SELFCONT   (1            ) 000           8TH GRADE ORGANIZATION:  SELF CONTAINED
002   SEMIDEPT   (2            ) 100           8TH GRADE ORGANIZATION:  SEMI-DEPARTMENTALIZED
003   DEPTLIZD   (3            ) 010           8TH GRADE ORGANIZATION:  DEPARTMENTALIZED
004   ORGANIZ?   (M            ) 001           8TH GRADE ORGANIZATION:  MISSING

CONDITIONING ID:            SCHL0009
DESCRIPTION:                ARE 4TH GRADERS ASSIGNED TO CLASSES BY ABILITY?
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                CLASS/AB            LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                    C031100             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:  3
001   ABILTY-Y   (1            ) 00            4TH GRADERS ASSIGNED BY ABILITY:  YES
002   ABILTY-N   (2            ) 10            4TH GRADERS ASSIGNED BY ABILITY:  NO
003   ABILTY-?   (M            ) 01            4TH GRADERS ASSIGNED BY ABILITY:  MISSING

CONDITIONING ID:            SCHL0010
DESCRIPTION:                ARE 8TH GRADERS ASSIGNED TO ENGLISH BY ABILITY?
GRADES/ASSESSMENTS:         N08, S08
GROUP LABEL:                ENG/AB8             LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                    C034401             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:  3
001   ENG/AB-Y   (1            ) 00            ENGLISH BY ABILITY:  YES
002   ENG/AB-N   (2            ) 10            ENGLISH BY ABILITY:  NO
003   ENG/AB-?   (M            ) 01            ENGLISH BY ABILITY:  MISSING

CONDITIONING ID:            SCHL0011
DESCRIPTION:                ARE 12TH GRADERS ASSIGNED TO ENGLISH BY ABILITY?
GRADES/ASSESSMENTS:         N12
GROUP LABEL:                ENG/AB12            LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                    C035001             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:  3
001   ENG/AB-Y   (1            ) 00            ENGLISH BY ABILITY:  YES
002   ENG/AB-N   (2            ) 10            ENGLISH BY ABILITY:  NO
003   ENG/AB-?   (M            ) 01            ENGLISH BY ABILITY:  MISSING

CONDITIONING ID:            SCHL0012
DESCRIPTION:                NUMBER OF SEMESTERS ENGLISH/LITERATURE/WRITING REQUIRED (GRADE 12)
GRADES/ASSESSMENTS:         N12
GROUP LABEL:                S_SEMENG            LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                    C035201             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:  3
001   #SEMENG0   (1-8          ) 00            SEMESTERS ENGLISH:  NONE-SEVEN
002   #SEMENG8   (9            ) 10            SEMESTERS ENGLISH:  EIGHT
003   #SEMENG?   (M            ) 01            SEMESTERS ENGLISH:  MISSING

CONDITIONING ID:            SCHL0013
DESCRIPTION:                ARE 12TH GRADERS ASSIGNED TO MATHEMATICS BY ABILITY?
GRADES/ASSESSMENTS:         N12
GROUP LABEL:                MAT/AB12            LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                    C035002             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:  3
001   MAT/AB-Y   (1            ) 00            MATH BY ABILITY:  YES
002   MAT/AB-N   (2            ) 10            MATH BY ABILITY:  NO
003   MAT/AB-?   (M            ) 01            MATH BY ABILITY:  MISSING

CONDITIONING ID:            SCHL0014
DESCRIPTION:                POLICY CONTROLLING TIME FOR READING INSTRUCTION?
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                POLICY-R            LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                    C031301             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:  3
001   RD_POL-Y   (1            ) 00            READING TIME POLICY:  YES
```

BEST COPY AVAILABLE

Table F-2 (continued)
Conditioning Variables for Reading Main Samples

```
002   RD_POL-N   (2            )  10          READING TIME POLICY:  NO
003   RD_POL-?   (M            )  01          READING TIME POLICY:  MISSING

CONDITIONING ID:             SCHL0015
DESCRIPTION:                 POLICY CONTROLLING TIME FOR WRITING INSTRUCTION?
GRADES/ASSESSMENTS:          N04, S04
GROUP LABEL:                 POLICY-W         LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                     C031302          DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS            NUMBER OF SPECIFICATION RECORDS:  3
001   WR_POL-Y   (1            )  00          WRITING TIME POLICY:  YES
002   WR_POL-N   (2            )  10          WRITING TIME POLICY:  NO
003   WR_POL-?   (M            )  01          WRITING TIME POLICY:  MISSING

CONDITIONING ID:             SCHL0016
DESCRIPTION:                 ARE COMPUTERS ALWAYS AVAILABLE IN ENGLISH CLASSROOMS? (GRADE 8)
GRADES/ASSESSMENTS:          N08, S08, N12
GROUP LABEL:                 COMP_ECL         LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                     C035601          DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS            NUMBER OF SPECIFICATION RECORDS:  3
001   CMPECL-Y   (1            )  00          COMPUTERS AVAILABLE IN ENGLISH CLASS:  YES
002   CMPECL-N   (2            )  10          COMPUTERS AVIALABLE IN ENGLISH CLASS:  NO
003   CMPECL-?   (M            )  01          COMPUTERS AVIALABLE IN ENGLISH CLASS:  MISSING

CONDITIONING ID:             SCHL0017
DESCRIPTION:                 ARE COMPUTERS ALWAYS GROUPED IN A LAB AVAILABLE FOR ENGLISH CLASSES? (GRADE 8)
GRADES/ASSESSMENTS:          N08, S08, N12
GROUP LABEL:                 COMP_ELB         LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                     C035602          DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS            NUMBER OF SPECIFICATION RECORDS:  3
001   CMPELB-Y   (1            )  00          COMPUTERS IN LAB FOR ENGLISH CLASS:  YES
002   CMPELB-N   (2            )  10          COMPUTERS IN LAB FOR ENGLISH CLASS:  NO
003   CMPELB-?   (M            )  01          COMPUTERS IN LAB FOR ENGLISH CLASS:  MISSING

CONDITIONING ID:             SCHL0018
DESCRIPTION:                 ARE COMPUTERS AVAILABLE TO BRING TO ENGLISH CLASSES? (GRADE 8)
GRADES/ASSESSMENTS:          N08, S08, N12
GROUP LABEL:                 COMP_EBR         LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                     C035603          DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS            NUMBER OF SPECIFICATION RECORDS:  3
001   CMPEBR-Y   (1            )  00          COMPUTERS AVAILABLE IN ENGLISH CLASS:  YES
002   CMPEBR-N   (2            )  10          COMPUTERS AVIALABLE IN ENGLISH CLASS:  NO
003   CMPEBR-?   (M            )  01          COMPUTERS AVIALABLE IN ENGLISH CLASS:  MISSING

CONDITIONING ID:             SCHL0019
DESCRIPTION:                 NUMBER OF STUDENTS ENROLLED IN AP ENGLISH
GRADES/ASSESSMENTS:          N12
GROUP LABEL:                 #AP_ENGL         LENGTH OF CONTRAST FIELD       :  6
NAEP ID:                     C035802          DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS            NUMBER OF SPECIFICATION RECORDS:  7
001   AP_ENGL1   (1            )  000000      # STUDENTS AP ENGLISH:  NONE
002   AP_ENGL2   (2            )  100000      # STUDENTS AP ENGLISH:  1-5
003   AP_ENGL3   (3            )  010000      # STUDENTS AP ENGLISH:  6-10
004   AP_ENGL4   (4            )  001000      # STUDENTS AP ENGLISH:  11-25
005   AP_ENGL5   (5            )  000100      # STUDENTS AP ENGLISH:  26-50
006   AP_ENGL6   (6            )  000010      # STUDENTS AP ENGLISH:  MORE THAN 50
007   AP_ENGL?   (M            )  000001      # STUDENTS AP ENGLISH:  MISSING

CONDITIONING ID:             SCHL0020
DESCRIPTION:                 DOES SCHOOL INVOLVE PARENTS AS AIDES IN CLASS?
GRADES/ASSESSMENTS:          N04, S04, N08, S08, N12
GROUP LABEL:                 PAR_AIDE         LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                     C032207          DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS            NUMBER OF SPECIFICATION RECORDS:  4
001   PARAID-R   (1            )  000         PARENTS AS AIDES IN CLASS:  ROUTINELY
002   PARAID-O   (2            )  100         PARENTS AS AIDES IN CLASS:  OCCASIONALLY
```

Table F-2 (continued)
Conditioning Variables for Reading Main Samples

```
003   PARAID-N   (3           )  010         PARENTS AS AIDES IN CLASS:  NO
004   PARAID-?   (M           )  001         PARENTS AS AIDES IN CLASS:  MISSING

CONDITIONING ID:            SCHL0021
DESCRIPTION:                NUMBER OF MINUTES OF DAILY ENGLISH INSTRUCTIONAL PERIOD
GRADES/ASSESSMENTS:         N12
GROUP LABEL:                MIN/ENGL            LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                    C035900             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   5
001   ENGL<30M   (1           )  0000        ENGLISH PERIOD LENGTH:  30 MINTUES OR LESS
002   ENGL=45M   (2           )  1000        ENGLISH PERIOD LENGTH:  45 MINUTES
003   ENGL=60M   (3           )  0100        ENGLISH PERIOD LENGTH:  60 MINUTES
004   ENGL>90M   (4           )  0010        ENGLISH PERIOD LENGTH:  90 MINUTES OR MORE
005   MIN/ENG?   (M           )  0001        ENGLISH PERIOD LENGTH:  MISSING

CONDITIONING ID:            TCHR0001
DESCRIPTION:                HOW WELL DOES SCHOOL PROVIDE RESOURCES
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                RESOURCE            LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                    T041201             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   5
001   RESOURC1   (1           )  0000        RESOURCES:  GET ALL
002   RESOURC2   (2           )  1000        RESOURCES:  GET MOST
003   RESOURC3   (3           )  0100        RESOURCES:  GET SOME
004   RESOURC4   (4           )  0010        RESOURCES:  DON'T GET
005   RESOURC?   (M,DNA       )  0001        RESOURCES:  MISSING, DOES NOT APPLY

CONDITIONING ID:            TCHR0002
DESCRIPTION:                TEACHER MATCH STATUS WITH STUDENT
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                T_MATCH             LENGTH OF CONTRAST FIELD       :   2
NAEP ID:                    TCHMTCH             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   3
001   TMCH-NO    (1,M         )  00          TEACHER MATCH:  NO MATCH
002   TMCH-PAR   (2           )  10          TEACHER MATCH:  PARTIAL MATCH
003   TMCH-COM   (3           )  01          TEACHER MATCH:  COMPLETE MATCH

CONDITIONING ID:            TCHR0003
DESCRIPTION:                TEACHER GENDER
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                T_GENDER            LENGTH OF CONTRAST FIELD       :   2
NAEP ID:                    T040001             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   3
001   T_MALE     (1           )  00          TEACHER GENDER:  MALE
002   T_FEMALE   (2           )  10          TEACHER GENDER:  FEMALE
003   T_SEX-?    (M,DNA       )  01          TEACHER GENDER:  MISSING, DOES NOT APPLY

CONDITIONING ID:            TCHR0004
DESCRIPTION:                TEACHER RACE/ETHNICITY
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                T_RACE              LENGTH OF CONTRAST FIELD       :   5
NAEP ID:                    T040101             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   6
001   T_WHITE    (1           )  00000       TEACHER ETHNICITY:  WHITE
002   T_BLACK    (2           )  10000       TEACHER ETHNICITY:  BLACK
003   T_HISP     (3           )  01000       TEACHER ETHNICITY:  HISPANIC
004   T_ASIAN    (4           )  00100       TEACHER ETHNICITY:  ASIAN, PACIFIC ISLANDER
005   T_AM.IND   (5           )  00010       TEACHER ETHNICITY:  AMERICAN INDIAN, ALASKAN NATIVE
006   T_RACE-?   (M,DNA       )  00001       TEACHER ETHNICITY:  MISSING, DOES NOT APPLY

CONDITIONING ID:            TCHR0005
DESCRIPTION:                TEACHER HISPANIC BACKGROUND
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                T_HISPBK            LENGTH OF CONTRAST FIELD       :   5
NAEP ID:                    T040201             DEGREES OF FREEDOM PER CONTRAST:   1
```

Table F-2 (continued)
Conditioning Variables for Reading Main Samples

```
TYPE OF CONTRAST:            CLASS                NUMBER OF SPECIFICATION RECORDS:   6
001   T_NONHSP    (1             )   00000        TEACHER HISPANIC BACKGROUND:  NOT HISPANIC
002   T_MEXICN    (2             )   10000        TEACHER HISPANIC BACKGROUND:  MEXICAN/MEXICAN AMERICAN
003   T_PUERTO    (3             )   01000        TEACHER HISPANIC BACKGROUND:  PUERTO RICAN
004   T_CUBAN     (4             )   00100        TEACHER HISPANIC BACKGROUND:  CUBAN
005   T_OTHER     (5             )   00010        TEACHER HISPANIC BACKGROUND:  OTHER
006   T_HISP-?    (M,DNA         )   00001        TEACHER HISPANIC BACKGROUND:  MISSING, DOES NOT APPLY

CONDITIONING ID:             TCHR0006
DESCRIPTION:                 YEARS TEACHING ELEMENTARY/SECONDARY SCHOOL
GRADES/ASSESSMENTS:          N04, S04
GROUP LABEL:                 T_YRSEXP             LENGTH OF CONTRAST FIELD       :   5
NAEP ID:                     T040301              DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:            CLASS                NUMBER OF SPECIFICATION RECORDS:   6
001   T_YREXP1    (1             )   00000        YEARS TEACHING:  2 OR LESS YEARS
002   T_YREXP2    (2             )   10000        YEARS TEACHING:  3-5 YEARS
003   T_YREXP3    (3             )   01000        YEARS TEACHING:  6-10 YEARS
004   T_YREXP4    (4             )   00100        YEARS TEACHING:  11-24 YEARS
005   T_YREXP5    (5             )   00010        YEARS TEACHING:  25 OR MORE YEARS
006   T_YREXP?    (M,DNA         )   00001        YEARS TEACHING:  MISSING, DOES NOT APPLY

CONDITIONING ID:             TCHR0007
DESCRIPTION:                 TYPE OF TEACHING CERTIFICATION
GRADES/ASSESSMENTS:          N04, S04
GROUP LABEL:                 TCH_CERT             LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                     T040401              DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:            CLASS                NUMBER OF SPECIFICATION RECORDS:   4
001   TCERT-NO    (1             )   000          TEACHING CERTIFICATION:  NONE, TEMPORARY, PROVISIONAL
002   TCERT-RG    (2             )   100          TEACHING CERTIFICATION:  REGULAR, NOT HIGHEST AVAILABLE
003   TCERT-HI    (3             )   010          TEACHING CERTIFICATION:  HIGHEST AVAILABLE
004   TCERT-?     (M,DNA         )   001          TEACHING CERTIFICATION:  MISSING, DOES NOT APPLY

CONDITIONING ID:             TCHR0008
DESCRIPTION:                 TEACHER GENERAL CERTIFICATION (ELEMENTARY, MIDDLE/JUNIOR HS EDUCATION)
GRADES/ASSESSMENTS:          N04, S04
GROUP LABEL:                 CERT-GEN             LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                     T040501              DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:            CLASS                NUMBER OF SPECIFICATION RECORDS:   4
001   CERTG-Y     (1             )   000          GENERAL CERTIFICATION:  YES
002   CERTG-N     (2             )   100          GENERAL CERTIFICATION:  NO
003   CERTG-NS    (3             )   010          GENERAL CERTIFICATION:  NOT OFFERED IN STATE
004   CERTG-?     (M,DNA         )   001          GENERAL CERTIFICATION:  MISSING, DOES NOT APPLY

CONDITIONING ID:             TCHR0009
DESCRIPTION:                 TEACHER'S HIGHEST ACADEMIC DEGREE
GRADES/ASSESSMENTS:          N04, S04
GROUP LABEL:                 T_DEGREE             LENGTH OF CONTRAST FIELD       :   6
NAEP ID:                     T040601              DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:            CLASS                NUMBER OF SPECIFICATION RECORDS:   7
001   <BACHLRS    (1             )   000000       TEACHER DEGREE:  LESS THAN A BACHELOR'S DEGREE
002   BACHELRS    (2             )   100000       TEACHER DEGREE:  BACHELOR'S DEGREE
003   MASTERS     (3             )   010000       TEACHER DEGREE:  MASTER'S DEGREE
004   SPECLIST    (4             )   001000       TEACHER DEGREE:  EDUCATION SPECIALIST
005   DOCTORAT    (5             )   000100       TEACHER DEGREE:  DOCTORATE
006   PROFESSL    (6             )   000010       TEACHER DEGREE:  PROFESSIONAL DEGREE
007   DEGREE-?    (M,DNA         )   000001       TEACHER DEGREE:  MISSING, DOES NOT APPLY

CONDITIONING ID:             TCHR0010
DESCRIPTION:                 TEACHER UNDERGRADUATE MAJOR IN EDUCATION
GRADES/ASSESSMENTS:          N04, S04
GROUP LABEL:                 UGRAD_ED             LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                     T040701              DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:            CLASS                NUMBER OF SPECIFICATION RECORDS:   2
001   UGR_ED-?    (M,DNA         )   0            UNDERGRADUATE EDUCATION MAJOR:  MISSING, DOES NOT APPLY
002   UGR_ED-Y    (1             )   1            UNDERGRADUATE EDUCATION MAJOR:  YES
```

Table F-2 (continued)
Conditioning Variables for Reading Main Samples

```
CONDITIONING ID:          TCHR0011
DESCRIPTION:              TEACHER UNDERGRADUATE MAJOR IN ENGLISH/READING/LANGUAGE ARTS
GRADES/ASSESSMENTS:       N04, S04
GROUP LABEL:              UGRAD_RD            LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                  T040702             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS               NUMBER OF SPECIFICATION RECORDS:  2
001   UGR_RD-?   (M,DNA     ) 0               UNDERGRADUATE READING MAJOR:  MISSING, DOES NOT APPLY
002   UGR_RD-Y   (1         ) 1               UNDERGRADUATE READING MAJOR:  YES

CONDITIONING ID:          TCHR0012
DESCRIPTION:              TEACHER GRADUATE MAJOR IN EDUCATION
GRADES/ASSESSMENTS:       N04, S04
GROUP LABEL:              GRAD_ED             LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                  T040801             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS               NUMBER OF SPECIFICATION RECORDS:  2
001   GR_ED-?    (M,DNA     ) 0               GRADUATE EDUCATION MAJOR:  MISSING, DOES NOT APPLY
002   GR_ED-Y    (1         ) 1               GRADUATE EDUCATION MAJOR:  YES

CONDITIONING ID:          TCHR0013
DESCRIPTION:              TEACHER GRADUATE MAJOR IN ENGLISH/READING/LANGUAGE ARTS
GRADES/ASSESSMENTS:       N04, S04
GROUP LABEL:              GRAD_RD             LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                  T040802             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS               NUMBER OF SPECIFICATION RECORDS:  2
001   GR_RD-?    (M,DNA     ) 0               GRADUATE READING MAJOR:  MISSING, DOES NOT APPLY
002   GR_RD-Y    (1         ) 1               GRADUATE READING MAJOR:  YES

CONDITIONING ID:          TCHR0014
DESCRIPTION:              NO TEACHER GRADUATE-LEVEL STUDY
GRADES/ASSESSMENTS:       N04, S04
GROUP LABEL:              GRAD_NO             LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                  T040806             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS               NUMBER OF SPECIFICATION RECORDS:  2
001   GR_NO-?    (M,DNA     ) 0               NO GRAUDATE STUDY:  MISSING, DOES NOT APPLY
002   GR_NO-Y    (1         ) 1               NO GRADUATE STUDY:  YES

CONDITIONING ID:          TCHR0015
DESCRIPTION:              ARE CURRICULUM SPECIALISTS AVAILABLE FOR READING?
GRADES/ASSESSMENTS:       N04, S04
GROUP LABEL:              CURSPE-R            LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                  T041301             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS               NUMBER OF SPECIFICATION RECORDS:  3
001   READCS-Y   (1         ) 00              READING CURRICULUM SPECIALISTS:  YES
002   READCS-N   (2         ) 10              READING CURRICULUM SPECIALISTS:  NO
003   READCS-?   (M,DNA     ) 01              READING CURRICULUM SPECIALISTS:  MISSING, DOES NOT APPLY

CONDITIONING ID:          TCHR0016
DESCRIPTION:              HOW OFTEN DO AIDES ASSIST YOU IN CLASS?
GRADES/ASSESSMENTS:       N04, S04
GROUP LABEL:              TCH_AIDE            LENGTH OF CONTRAST FIELD       :  5
NAEP ID:                  T041401             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS               NUMBER OF SPECIFICATION RECORDS:  6
001   TCHAIDE1   (1         ) 00000           TEACHER AIDES IN CLASS:  EVERY DAY
002   TCHAIDE2   (2         ) 10000           TEACHER AIDES IN CLASS:  SEVERAL TIMES A WEEK
003   TCHAIDE3   (3         ) 01000           TEACHER AIDES IN CLASS:  ONCE A WEEK
004   TCHAIDE4   (4         ) 00100           TEACHER AIDES IN CLASS:  LESS THAN ONCE A WEEK
005   TCHAIDE5   (5         ) 00010           TEACHER AIDES IN CLASS:  NEVER
006   TCHAIDE?   (M,DNA     ) 00001           TEACHER AIDES IN CLASS:  MISSING, DOES NOT APPLY

CONDITIONING ID:          TCHR0017
DESCRIPTION:              NUMBER OF STUDENTS IN CLASS
GRADES/ASSESSMENTS:       N04, S04
GROUP LABEL:              T_NCLASS            LENGTH OF CONTRAST FIELD       :  5
NAEP ID:                  TCHNCLS             DEGREES OF FREEDOM PER CONTRAST:  1
```

Table F-2 (continued)
Conditioning Variables for Reading Main Samples

```
TYPE OF CONTRAST:          CLASS                NUMBER OF SPECIFICATION RECORDS:   6
001   T_NCLAS1   (0-20       )  00000       CLASS SIZE:  0-20
002   T_NCLAS2   (21-25      )  10000       CLASS SIZE:  21-25
003   T_NCLAS3   (26-31      )  01000       CLASS SIZE:  26-30
004   T_NCLAS4   (31-35      )  00100       CLASS SIZE:  31-35
005   T_NCLAS5   (36-61      )  00010       CLASS SIZE:  36-60
006   T_NCLAS?   (M          )  00001       CLASS SIZE:  MISSING

CONDITIONING ID:           TRED0001
DESCRIPTION:               TEACHER HOURS SPENT IN IN-SERVICE READING EDUCATION
GRADES/ASSESSMENTS:        N04, S04
GROUP LABEL:               INSERV_R             LENGTH OF CONTRAST FIELD       :   5
NAEP ID:                   T041001              DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS                NUMBER OF SPECIFICATION RECORDS:   6
001    INSERVR1   (1         )  00000       READING HOURS IN-SERVICE:  NONE
002    INSERVR2   (2         )  10000       READING HOURS IN-SERVICE:  LESS THAN 6 HOURS
003    INSERVR3   (3         )  01000       READING HOURS IN-SERVICE:  6-15 HOURS
004    INSERVR4   (4         )  00100       READING HOURS IN-SERVICE:  16-35 HOURS
005    INSERVR5   (5         )  00010       READING HOURS IN-SERVICE:  MORE THAN 35 HOURS
006    INSERVR?   (M,DNA     )  00001       READING HOURS IN-SERVICE:  MISSING, DOES NOT APPLY

CONDITIONING ID:           TRED0002
DESCRIPTION:               TEACHER CERTIFICATION IN READING
GRADES/ASSESSMENTS:        N04, S04
GROUP LABEL:               CERT-RED             LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                   T040502              DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS                NUMBER OF SPECIFICATION RECORDS:   4
001    CERTR-Y    (1         )  000         READING CERTIFICATION:  YES
002    CERTR-N    (2         )  100         READING CERTIFICATION:  NO
003    CERTR-NS   (3         )  010         READING CERTIFICATION:  NOT OFFERED IN STATE
004    CERTR-?    (M,DNA     )  001         READING CERTIFICATION:  MISSING, DOES NOT APPLY

CONDITIONING ID:           TRED0003
DESCRIPTION:               TEACHER CERTIFICATION MIDDLE/JUNIOR HS/SECONDARY ENGLISH/LANGUAGE ARTS
GRADES/ASSESSMENTS:        N04, S04
GROUP LABEL:               CERT-ENG             LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                   T040503              DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS                NUMBER OF SPECIFICATION RECORDS:   4
001    CERTE-Y    (1         )  000         READING CERTIFICATION:  YES
002    CERTE-N    (2         )  100         READING CERTIFICATION:  NO
003    CERTE-NS   (3         )  010         READING CERTIFICATION:  NOT OFFERED IN STATE
004    CERTE-?    (M,DNA     )  001         READING CERTIFICATION:  MISSING, DOES NOT APPLY

CONDITIONING ID:           TRED0004
DESCRIPTION:               ARE STUDENTS ASSIGNED TO READING CLASS BY ABILITY?
GRADES/ASSESSMENTS:        N04, S04
GROUP LABEL:               ABIL_CLA             LENGTH OF CONTRAST FIELD       :   2
NAEP ID:                   T046101              DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS                NUMBER OF SPECIFICATION RECORDS:   3
001    AB_CLA-Y   (1         )  00          READING CLASS BY ABILITY: YES
002    AB_CLA-N   (2         )  10          READING CLASS BY ABILITY: NO
003    AB_CLA-?   (M,DNA     )  01          READING CLASS BY ABILITY: MISSING, DOES NOT APPLY

CONDITIONING ID:           TRED0005
DESCRIPTION:               READING ABILITY LEVEL OF STUDENTS IN CLASS
GRADES/ASSESSMENTS:        N04, S04
GROUP LABEL:               ABIL_RED             LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                   T046201              DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS                NUMBER OF SPECIFICATION RECORDS:   5
001    AB_READ1   (1         )  0000        READING ABILITY:  PRIMARILY HIGH ABILITY
002    AB_READ2   (2         )  1000        READING ABILITY:  PRIMARILY AVERAGE ABILITY
003    AB_READ3   (3         )  0100        READING ABILITY:  PRIMARILY LOW ABILITY
004    AB_READ4   (4         )  0010        READING ABILITY:  WIDELY MIXED ABILITY
005    AB_READ?   (M,DNA     )  0001        READING ABILITY:  MISSING, DOES NOT APPLY
```

Table F-2 (continued)
Conditioning Variables for Reading Main Samples

```
CONDITIONING ID:            TRED0006
DESCRIPTION:                TIME SPENT PER DAY ON READING INSTRUCTION
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                INS_TIME           LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                    T046301            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS              NUMBER OF SPECIFICATION RECORDS:  4
001   INSTIME1   (1,2          )  000          READING INSTRUCTION TIME:  30-45 MINUTES/DAY
002   INSTIME2   (3            )  100          READING INSTRUCTION TIME:  60 MINUTES/DAY
003   INSTIME3   (4            )  010          READING INSTRUCTION TIME:  90 MINUTES OR MORE/DAY
004   INSTIME?   (M,DNA        )  001          READING INSTRUCTION TIME:  MISSING, DOES NOT APPLY

CONDITIONING ID:            TRED0007
DESCRIPTION:                NUMBER OF INSTRUCTIONAL GROUPS CLASS DIVIDED INTO FOR READING
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                #INS_GRP           LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                    T046400            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS              NUMBER OF SPECIFICATION RECORDS:  4
001   INSGRP-1   (1            )  000          NUMBER INSTRUCTIONAL GROUPS:  WHOLE CLASS ACTIVITY
002   INSGRP-2   (2            )  100          NUMBER INSTRUCTIONAL GROUPS:  WHOLE CLASS/FLEXIBLE GROUPING
003   INSGRP-3   (3,4,5,6      )  010          NUMBER INSTRUCTIONAL GROUPS:  TWO OR MORE
004   INSGRP-4   (7,M,DNA      )  001          NUMBER INSTR GROUPS:  INDIV INSTRUCTION, MISSING, NOT APP

CONDITIONING ID:            TRED0008
DESCRIPTION:                TYPE OF MATERIALS FORMING CORE OF READING PROGRAM
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                RMATERLS           LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                    T046501            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS              NUMBER OF SPECIFICATION RECORDS:  5
001   BASAL_RM   (1            )  0000         TYPE OF READING MATERIALS:  BASAL
002   TRADE_RM   (2            )  1000         TYPE OF READING MATERIALS:  TRADE
003   BAS&TRA    (3            )  0100         TYPE OF READING MATERIALS:  BASAL AND TRADE
004   OTHER_RM   (4            )  0010         TYPE OF READING MATERIALS:  OTHER
005   RMATS-?    (M,DNA        )  0001         TYPE OF READING MATERIALS:  MISSING, DOES NOT APPLY

CONDITIONING ID:            TRED0009
DESCRIPTION:                HOW OFTEN ARE NEWSPAPERS/MAGAZINES USED TO TEACH READING CLASS
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                NEWS/MAG           LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                    T046601            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS              NUMBER OF SPECIFICATION RECORDS:  4
001   NEWMAG-1   (1,2          )  000          NEWS/MAGS (TCHR):  ALMOST EVERY DAY, ONCE OR TWICE/WEEK
002   NEWMAG-2   (3            )  100          NEWSPAPERS/MAGAZINES (TEACHER):  ONCE OR TWICE A MONTH
003   NEWMAG-3   (4            )  010          NEWSPAPERS/MAGAZINES (TEACHER):  NEVER OF HARDLEY EVER
004   NEWMAG-?   (M,DNA        )  001          NEWSPAPERS/MAGAZINES (TEACHER):  MISSING, DOES NOT APPLY

CONDITIONING ID:            TRED0010
DESCRIPTION:                HOW OFTEN ARE READING KITS USED TO TEACH READING CLASS
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                READKITS           LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                    T046602            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS              NUMBER OF SPECIFICATION RECORDS:  4
001   RDKITS-1   (1,2          )  000          READING KITS (TCHR):  ALMOST EVERY DAY, ONCE OR TWICE A WEEK
002   RDKITS-2   (3            )  100          READING KITS (TEACHER):  ONCE OR TWICE A MONTH
003   RDKITS-3   (4            )  010          READING KITS (TEACHER):  NEVER OF HARDLEY EVER
004   RDKITS-?   (M,DNA        )  001          READING KITS (TEACHER):  MISSING, DOES NOT APPLY

CONDITIONING ID:            TRED0011
DESCRIPTION:                HOW OFTEN IS READING COMPUTER SOFTWARE USED TO TEACH READING CLASS
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                SOFTWARE           LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                    T046603            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS              NUMBER OF SPECIFICATION RECORDS:  4
001   SOFTWA-1   (1,2          )  000          READ COMP SOFT (TCHR):  ALMOST EVERY DAY, ONCE OR TWICE/WEEK
002   SOFTWA-2   (3            )  100          READING COMPUTER SOFTWARE (TEACHER):  ONCE OR TWICE A MONTH
003   SOFTWA-3   (4            )  010          READING COMPUTER SOFTWARE (TEACHER):  NEVER OF HARDLEY EVER
```

Table F-2 (continued)
Conditioning Variables for Reading Main Samples

```
004   SOFTWA-?   (M,DNA      )  001         READING COMPUTER SOFTWARE (TEACHER):  MISSING, NOT APPLY

CONDITIONING ID:            TRED0012
DESCRIPTION:                HOW OFTEN ARE VARIETY OF BOOKS USED TO TEACH READING CLASS
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                VARI_BKS            LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                    T046604             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   4
001   VARBKS-1   (1            )  000           VARIETY OF BOOKS:  ALMOST EVERY DAY
002   VARBKS-2   (2            )  100           VARIETY OF BOOKS:  ONCE OR TWICE A WEEK
003   VARBKS-3   (3,4          )  010           VARIETY OF BOOKS:  ONCE OR TWICE A MONTH, NEVER/HARDLY EVER
004   VARBKS-?   (M,DNA        )  001           VARIETY OF BOOKS:  MISSING, DOES NOT APPLY

CONDITIONING ID:            TRED0013
DESCRIPTION:                HOW OFTEN ARE MATERIALS FROM OTHER SUBJECTS USED TO TEACH READING CLASS
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                OTH_MATS            LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                    T046605             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   4
001   OTHMAT-1   (1            )  000           OTHER SUBJECT MATERIALS (TEACHER):  ALMOST EVERY DAY
002   OTHMAT-2   (2            )  100           OTHER SUBJECT MATERIALS (TEACHER):  ONCE OR TWICE A WEEK
003   OTHMAT-3   (3,4          )  010           OTHER SUB MATS (TCHR):ONCE OR TWICE/MONTH, NEVER/HARDLY EVER
004   OTHMAT-?   (M,DNA        )  001           OTHER SUBJECT MATERIALS (TEACHER):  MISSING, DOES NOT APPLY

CONDITIONING ID:            TRED0014
DESCRIPTION:                HOW DO YOU DISCUSS NEW OR DIFFICULT VOCABULARY?
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                T_VOCAB             LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                    T046701             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   5
001   VOCAB-T1   (1            )  0000          DISCUSS VOCABULARY:  ALMOST EVERY DAY
002   VOCAB-T2   (2            )  1000          DISCUSS VOCABULARY:  ONCE OR TWICE A WEEK
003   VOCAB-T3   (3            )  0100          DISCUSS VOCABULARY:  ONCE OR TWICE A MONTH
004   VOCAB-T4   (4            )  0010          DISCUSS VOCABULARY:  NEVER OR HARDLY EVER
005   VOCAB-T?   (M,DNA        )  0001          DISCUSS VOCABULARY:  MISSING, DOES NOT APPLY

CONDITIONING ID:            TRED0015
DESCRIPTION:                HOW OFTEN DO YOU ASK STUDENTS TO READ ALOUD?
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                T_ALOUD             LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                    T046702             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   5
001   ALOUD-T1   (1            )  0000          READ ALOUD: ALMOST EVERY DAY
002   ALOUD-T2   (2            )  1000          READ ALOUD: ONCE OR TWICE A WEEK
003   ALOUD-T3   (3            )  0100          READ ALOUD: ONCE OR TWICE A MONTH
004   ALOUD-T4   (4            )  0010          READ ALOUD: NEVER OR HARDLY EVER
005   ALOUD-T?   (M,DNA        )  0001          READ ALOUD: MISSING, DOES NOT APPLY

CONDITIONING ID:            TRED0016
DESCRIPTION:                HOW OFTEN DO YOU ASK STUDENTS TO TALK TO EACH OTHER ABOUT WHAT THEY HAVE READ?
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                T_TALKRD            LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                    T046703             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   5
001   TLKRD-T1   (1            )  0000          TALK ABOUT READING:  ALMOST EVERY DAY
002   TLKRD-T2   (2            )  1000          TALK ABOUT READING:  ONCE OR TWICE A WEEK
003   TLKRD-T3   (3            )  0100          TALK ABOUT READING:  ONCE OR TWICE A MONTH
004   TLKRD-T4   (4            )  0010          TALK ABOUT READING:  NEVER OR HARDLY EVER
005   TLKRD-T?   (M,DNA        )  0001          TALK ABOUT READING:  MISSING, DOES NOT APPLY

CONDITIONING ID:            TRED0017
DESCRIPTION:                HOW OFTEN DO YOU ASK STUDENTS TO WRITE SOMETHING ABOUT WHAT THEY HAVE READ?
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                T_WRITRD            LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                    T046704             DEGREES OF FREEDOM PER CONTRAST:   1
```

Table F-2 (continued)
Conditioning Variables for Reading Main Samples

```
TYPE OF CONTRAST:           CLASS                NUMBER OF SPECIFICATION RECORDS:   5
001   WRTRD-T1    (1           )  0000          WRITE ABOUT READING:  ALMOST EVERY DAY
002   WRTRD-T2    (2           )  1000          WRITE ABOUT READING:  ONCE OR TWICE A WEEK
003   WRTRD-T3    (3           )  0100          WRITE ABOUT READING:  ONCE OR TWICE A MONTH
004   WRTRD-T4    (4           )  0010          WRITE ABOUT READING:  NEVER OR HARDLY EVER
005   WRTRD-T?    (M,DNA       )  0001          WRITE ABOUT READING:  MISSING, DOES NOT APPLY

CONDITIONING ID:            TRED0018
DESCRIPTION:                HOW OFTEN DO YOU ASK STUDENTS TO WORK IN A READING WORKBOOK OR ON A WORKSHEET?
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                T_WBKWSH             LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                    T046705              DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS                NUMBER OF SPECIFICATION RECORDS:   5
001   WB/WS-T1    (1           )  0000          READING WORKBOOK/WORKSHEET:  ALMOST EVERY DAY
002   WB/WS-T2    (2           )  1000          READING WORKBOOK/WORKSHEET:  ONCE OR TWICE A WEEK
003   WB/WS-T3    (3           )  0100          READING WORKBOOK/WORKSHEET:  ONCE OR TWICE A MONTH
004   WB/WS-T4    (4           )  0010          READING WORKBOOK/WORKSHEET:  NEVER OR HARDLY EVER
005   WB/WS-T?    (M,DNA       )  0001          READING WORKBOOK/WORKSHEET:  MISSING, MISSING NOT APPLY

CONDITIONING ID:            TRED0019
DESCRIPTION:                HOW OFTEN DO YOU ASK STUDENTS TO READ SILENTLY?
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                T_SILENT             LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                    T046706              DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS                NUMBER OF SPECIFICATION RECORDS:   5
001   SILNT-T1    (1           )  0000          READ SILENTLY: ALMOST EVERY DAY
002   SILNT-T2    (2           )  1000          READ SILENTLY: ONCE OR TWICE A WEEK
003   SILNT-T3    (3           )  0100          READ SILENTLY: ONCE OR TWICE A MONTH
004   SILNT-T4    (4           )  0010          READ SILENTLY: NEVER OR HARDLY EVER
005   SILNT-T?    (M,DNA       )  0001          READ SILENTLY: MISSING, DOES NOT APPLY

CONDITIONING ID:            TRED0020
DESCRIPTION:                HOW OFTEN DO YOU GIVE STUDENTS TIME TO READ BOOKS OF THEIR OWN CHOOSING?
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                T_OWNBKS             LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                    T046707              DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS                NUMBER OF SPECIFICATION RECORDS:   5
001   OWNBK-T1    (1           )  0000          BOOKS CHOSEN YOURSELF:  ALMOST EVERY DAY
002   OWNBK-T2    (2           )  1000          BOOKS CHOSEN YOURSELF:  OR TWICE A WEEK
003   OWNBK-T3    (3           )  0100          BOOKS CHOSEN YOURSELF:  ONCE OR TWICE A MONTH
004   OWNBK-T4    (4           )  0010          BOOKS CHOSEN YOURSELF:  NEVER OR HARDLY EVER
005   OWNBK-T?    (M,DNA       )  0001          BOOKS CHOSEN YOURSELF:  MISSING, DOES NOT APPLY

CONDITIONING ID:            TRED0021
DESCRIPTION:                HOW OFTEN YOU ASK STUDENTS TO WRITE IN LOG OR JOURNAL ABOUT WHAT THEY HAVE READ?
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                T_RDLOG              LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                    T046708              DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS                NUMBER OF SPECIFICATION RECORDS:   5
001   RDLOG-T1    (1           )  0000          WRITE IN LOG/JOURNAL:  ALMOST EVERY DAY
002   RDLOG-T2    (2           )  1000          WRITE IN LOG/JOURNAL:  ONCE OR TWICE A WEEK
003   RDLOG-T3    (3           )  0100          WRITE IN LOG/JOURNAL:  ONCE OR TWICE A MONTH
004   RDLOG-T4    (4           )  0010          WRITE IN LOG/JOURNAL:  NEVER OR HARDLY EVER
005   RDLOG-T?    (M,DNA       )  0001          WRITE IN LOG/JOURNAL:  MISSING, DOES NOT APPLY

CONDITIONING ID:            TRED0022
DESCRIPTION:                HOW OFTEN YOU ASK STUDENTS TO DO GROUP ACTVTY/PROJECT ABOUT WHAT THEY HAVE READ?
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                T_RDPROJ             LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                    T046709              DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS                NUMBER OF SPECIFICATION RECORDS:   5
001   RDPRJ-T1    (1           )  0000          PROJECT ABOUT READING:  ALMOST EVERY DAY
002   RDPRJ-T2    (2           )  1000          PROJECT ABOUT READING:  ONCE OR TWICE A WEEK
003   RDPRJ-T3    (3           )  0100          PROJECT ABOUT READING:  ONCE OR TWICE A MONTH
004   RDPRJ-T4    (4           )  0010          PROJECT ABOUT READING:  NEVER OR HARDLY EVER
```

Table F-2 (continued)
Conditioning Variables for Reading Main Samples

```
005   RDPRJ-T?   (M,DNA       )  0001        PROJECT ABOUT READING:  MISSING, DOES NOT APPLY

CONDITIONING ID:            TRED0023
DESCRIPTION:                HOW MUCH READING INSTRUCTIONAL TIME DO YOU DEVOTE TO DECODING SKILLS?
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                DECODING             LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                    T046801              DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS                NUMBER OF SPECIFICATION RECORDS:  4
001   DECODE-1   (1           )  000         DECODING SKILLS:  ALMOST ALL OF THE TIME
002   DECODE-2   (2           )  100         DECODING SKILLS:  SOME OF THE TIME
003   DECODE-3   (3           )  010         DECODING SKILLS:  NEVER OF HARDLY EVER
004   DECODE-?   (M,DNA       )  001         DECODING SKILLS:  MISSING, DOES NOT APPLY

CONDITIONING ID:            TRED0024
DESCRIPTION:                HOW MUCH READING INSTRUCTIONAL TIME DO YOU DEVOTE TO ORAL READING?
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                ORALREAD             LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                    T046802              DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS                NUMBER OF SPECIFICATION RECORDS:  4
001   ORALRD-1   (1           )  000         ORAL READING:  ALMOST ALL OF THE TIME
002   ORALRD-2   (2           )  100         ORAL READING:  SOME OF THE TIME
003   ORALRD-3   (3           )  010         ORAL READING:  NEVER OF HARDLY EVER
004   ORALRD-?   (M,DNA       )  001         ORAL READING:  MISSING, DOES NOT APPLY

CONDITIONING ID:            TRED0025
DESCRIPTION:                HOW MUCH READING INSTRUCTIONAL TIME DO YOU DEVOTE TO VOCABULARY?
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                VOCABLRY             LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                    T046803              DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS                NUMBER OF SPECIFICATION RECORDS:  4
001   VOCABY-1   (1           )  000         VOCABULARY:  ALMOST ALL OF THE TIME
002   VOCABY-2   (2           )  100         VOCABULARY:  SOME OF THE TIME
003   VOCABY-3   (3           )  010         VOCABULARY:  NEVER OF HARDLY EVER
004   VOCABY-?   (M,DNA       )  001         VOCABULARY:  MISSING, DOES NOT APPLY

CONDITIONING ID:            TRED0026
DESCRIPTION:                HOW MUCH READING INSTRUCT TIME DO YOU DEVOTE TO COMPREHENSION/INTERPRETATION?
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                COMPREH              LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                    T046804              DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS                NUMBER OF SPECIFICATION RECORDS:  4
001   CMPREH-1   (1           )  000         COMPREHENSION/INTERPRETATION:  ALMOST ALL OF THE TIME
002   CMPREH-2   (2           )  100         COMPREHENSION/INTERPRETATION:  SOME OF THE TIME
003   CMPREH-3   (3           )  010         COMPREHENSION/INTERPRETATION:  NEVER OF HARDLY EVER
004   CMPREH-?   (M,DNA       )  001         COMPREHENSION/INTERPRETATION:  MISSING, DOES NOT APPLY

CONDITIONING ID:            TRED0027
DESCRIPTION:                HOW MUCH READING INSTRUCTIONAL TIME DO YOU DEVOTE TO READING STRATEGIES?
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                STRATGY              LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                    T046805              DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS                NUMBER OF SPECIFICATION RECORDS:  4
001   STRATG-1   (1           )  000         READING STRATEGIES:  ALMOST ALL OF THE TIME
002   STRATG-2   (2           )  100         READING STRATEGIES:  SOME OF THE TIME
003   STRATG-3   (3           )  010         READING STRATEGIES:  NEVER OF HARDLY EVER
004   STRATG-?   (M,DNA       )  001         READING STRATEGIES:  MISSING, DOES NOT APPLY

CONDITIONING ID:            TRED0028
DESCRIPTION:                HOW MUCH EMPHASIS DO YOU GIVE PHONICS?
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                EMP_PHON             LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                    T046901              DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS                NUMBER OF SPECIFICATION RECORDS:  4
001   EMPPHO-H   (1           )  000         EMPHASIS PHONICS:  HEAVY
002   EMPPHO-M   (2           )  100         EMPHASIS PHONICS:  MODERATE
```

Table F-2 (continued)
Conditioning Variables for Reading Main Samples

```
003   EMPPHO-L   (3          )  010         EMPHASIS PHONICS:  LITTLE OR NO
004   EMPPHO-?   (M,DNA      )  001         EMPHASIS PHONICS:  MISSING, DOES NOT APPLY

CONDITIONING ID:            TRED0029
DESCRIPTION:                HOW MUCH EMPHASIS DO YOU GIVE THE INTEGRATION OF READING AND WRITING?
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                EMP_INTG            LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                    T046902             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   EMPINT-H   (1          )  000         EMPHASIS INTEGRATION READING/WRITING:  HEAVY
002   EMPINT-M   (2          )  100         EMPHASIS INTEGRATION READING/WRITING:  MODERATE
003   EMPINT-L   (3          )  010         EMPHASIS INTEGRATION READING/WRITING:  LITTLE OR NO
004   EMPINT-?   (M,DNA      )  001         EMPHASIS INTEGRATION READING/WRITING:  MISSING, NOT APPLY

CONDITIONING ID:            TRED0030
DESCRIPTION:                HOW MUCH EMPHASIS DO YOU GIVE WHOLE LANGUAGE?
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                EMP_WLAN            LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                    T046903             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   EMPLAN-H   (1          )  000         EMPHASIS WHOLE LANGUAGE:  HEAVY
002   EMPLAN-M   (2          )  100         EMPHASIS WHOLE LANGUAGE:  MODERATE
003   EMPLAN-L   (3          )  010         EMPHASIS WHOLE LANGUAGE:  LITTLE OR NO
004   EMPLAN-?   (M,DNA      )  001         EMPHASIS WHOLE LANGUAGE:  MISSING, DOES NOT APPLY

CONDITIONING ID:            TRED0031
DESCRIPTION:                HOW MUCH EMPHASIS DO YOU GIVE LITERATURE-BASED READING?
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                EMP_LITB            LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                    T046904             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   EMPLIT-H   (1          )  000         EMPHASIS LITERATURE-BASED READING:  HEAVY
002   EMPLIT-M   (2          )  100         EMPHASIS LITERATURE-BASED READING:  MODERATE
003   EMPLIT-L   (3          )  010         EMPHASIS LITERATURE-BASED READING:  LITTLE OR NO
004   EMPLIT-?   (M,DNA      )  001         EMPHASIS LITERATURE-BASED READING:  MISSING, DOES NOT APPLY

CONDITIONING ID:            TRED0032
DESCRIPTION:                HOW MUCH EMPHASIS DO YOU GIVE READING ACROSS THE CONTENT AREAS?
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                EMP_CONT            LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                    T046905             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   EMPCON-H   (1          )  000         EMPHASIS READING ACROSS CONTENT AREAS:  HEAVY
002   EMPCON-M   (2          )  100         EMPHASIS READING ACROSS CONTENT AREAS:  MODERATE
003   EMPCON-L   (3          )  010         EMPHASIS READING ACROSS CONTENT AREAS:  LITTLE OR NO
004   EMPCON-?   (M,DNA      )  001         EMPHASIS READING ACROSS CONTENT AREAS:  MISSING, NOT APPLY

CONDITIONING ID:            TRED0033
DESCRIPTION:                HOW MUCH EMPHASIS DO YOU GIVE INDIVIDUALIZED READING PROGRAMS?
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                EMP_INDV            LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                    T046906             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   EMPCON-H   (1          )  000         EMPHASIS INDIVIDUALIZED READING PROGRAMS:  HEAVY
002   EMPCON-M   (2          )  100         EMPHASIS INDIVIDUALIZED READING PROGRAMS:  MODERATE
003   EMPCON-L   (3          )  010         EMPHASIS INDIVIDUALIZED READING PROGRAMS:  LITTLE OR NO
004   EMPCON-?   (M,DNA      )  001         EMPHASIS INDIVID READING PROGRAMS:  MISSING, NOT APPLY

CONDITIONING ID:            TRED0034
DESCRIPTION:                HOW OFTEN DO YOU USE MULTIPLE-CHOICE TESTS TO ASSESS STUDENTS IN READING?
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                MC_TESTS            LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                    T047001             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   MCTEST-1   (1          )  000         MULTIPLE-CHOICE TESTS:  ONCE OR TWICE A WEEK
```

Table F-2 (continued)
Conditioning Variables for Reading Main Samples

```
002   MCTEST-2   (2          )  100        MULTIPLE-CHOICE TESTS:  ONCE OR TWICE A MONTH
003   MCTEST-3   (3,4        )  010        MULT-CHOICE TESTS:  ONCE OR TWICE/YEAR, NEVER OR HARDLY EVER
004   MCTEST-?   (M,DNA      )  001        MULTIPLE-CHOICE TESTS:  MISSING, DOES NOT APPLY

CONDITIONING ID:             TRED0035
DESCRIPTION:                 HOW OFTEN DO YOU USE MULTIPLE-CHOICE TESTS TO ASSESS STUDENTS IN READING?
GRADES/ASSESSMENTS:          N04, S04
GROUP LABEL:                 SA_TESTS          LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                     T047002           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:            CLASS             NUMBER OF SPECIFICATION RECORDS:   4
001   SATEST-1   (1          )  000        SHORT-ANSWER TESTS:  ONCE OR TWICE A WEEK
002   SATEST-2   (2          )  100        SHORT-ANSWER TESTS:  ONCE OR TWICE A MONTH
003   SATEST-3   (3,4        )  010        SHORT-ANSWER TESTS:  ONCE OR TWICE/YEAR, NEVER/HARDLY EVER
004   SATEST-?   (M,DNA      )  001        SHORT-ANSWER TESTS:  MISSING, DOES NOT APPLY

CONDITIONING ID:             TRED0036
DESCRIPTION:                 HOW OFTEN DO YOU USE WRITING PARAGRAPHS TO ASSESS STUDENTS IN READING?
GRADES/ASSESSMENTS:          N04, S04
GROUP LABEL:                 WRT_PARA          LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                     T047003           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:            CLASS             NUMBER OF SPECIFICATION RECORDS:   4
001   WRTPAR-1   (1          )  000        ASSESS BY WRITING PARAGRAPHS:  ONCE OR TWICE A WEEK
002   WRTPAR-2   (2          )  100        ASSESS BY WRITING PARAGRAPHS:  ONCE OR TWICE A MONTH
003   WRTPAR-3   (3,4        )  010        ASSESS BY WRT PRGRPHS: ONCE OR TWICE/YEAR, NEVER/HARDLY EVER
004   WRTPAR-?   (M,DNA      )  001        ASSESS BY WRITING PARAGRAPHS:  MISSING, DOES NOT APPLY

CONDITIONING ID:             TRED0037
DESCRIPTION:                 HOW OFTEN DO YOU USE OBSERVATIONS TO ASSESS STUDENTS IN READING?
GRADES/ASSESSMENTS:          N04, S04
GROUP LABEL:                 OBSERVTN          LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                     T047004           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:            CLASS             NUMBER OF SPECIFICATION RECORDS:   4
001   OBSERV-1   (1          )  000        ASSESS BY OBSERVATIONS:  ONCE OR TWICE A WEEK
002   OBSERV-2   (2          )  100        ASSESS BY OBSERVATIONS:  ONCE OR TWICE A MONTH
003   OBSERV-3   (3,4        )  010        ASSESS BY OBS:  ONCE OR TWICE A YEAR, NEVER OR HARDLY EVER
004   OBSERV-?   (M,DNA      )  001        ASSESS BY OBSERVATIONS:  MISSING, DOES NOT APPLY

CONDITIONING ID:             TRED0038
DESCRIPTION:                 HOW OFTEN DO YOU USE ORAL READING TO ASSESS STUDENTS IN READING?
GRADES/ASSESSMENTS:          N04, S04
GROUP LABEL:                 ORALTEST          LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                     T047005           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:            CLASS             NUMBER OF SPECIFICATION RECORDS:   4
001   ORLTST-1   (1          )  000        ASSESS BY ORAL READING:  ONCE OR TWICE A WEEK
002   ORLTST-2   (2          )  100        ASSESS BY ORAL READING:  ONCE OR TWICE A MONTH
003   ORLTST-3   (3,4        )  010        ASSESS BY ORAL RDG:  ONCE OR TWICE/YEAR, NEVER/HARDLY EVER
004   ORLTST-?   (M,DNA      )  001        ASSESS BY ORAL READING:  MISSING, DOES NOT APPLY

CONDITIONING ID:             TRED0039
DESCRIPTION:                 HOW OFTEN USE INDIVID OR GROUP PROJECTS TO ASSESS STUDENTS IN READING?
GRADES/ASSESSMENTS:          N04, S04
GROUP LABEL:                 PROJECTS          LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                     T047006           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:            CLASS             NUMBER OF SPECIFICATION RECORDS:   4
001   PROJCT-1   (1          )  000        ASSESS BY PROJECTS/PRESENTATIONS:  ONCE OR TWICE A WEEK
002   PROJCT-2   (2          )  100        ASSESS BY PROJECTS/PRESENTATIONS:  ONCE OR TWICE A MONTH
003   PROJCT-3   (3,4        )  010        ASSESS BY PROJECTS/PRES:  ONCE OR TWICE/YEAR, NEVER/HARDLY
004   PROJCT-?   (M,DNA      )  001        ASSESS BY PROJECTS/PRESENTATIONS:  MISSING, DOES NOT APPLY

CONDITIONING ID:             TRED0040
DESCRIPTION:                 HOW OFTEN DO YOU USE READING PORTFOLIOS TO ASSESS STUDENTS IN READING?
GRADES/ASSESSMENTS:          N04, S04
GROUP LABEL:                 RD_PORTF          LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                     T047007           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:            CLASS             NUMBER OF SPECIFICATION RECORDS:   4
```

Table F-2 (continued)
Conditioning Variables for Reading Main Samples

```
001   RDPORT-1   (1            )  000         ASSESS BY READING PORTFOLIOS:  ONCE OR TWICE A WEEK
002   RDPORT-2   (2            )  100         ASSESS BY READING PORTFOLIOS:  ONCE OR TWICE A MONTH
003   RDPORT-3   (3,4          )  010         ASSESS BY RDG PORTFOLIOS:  ONCE OR TWICE/YEAR, NEVER/HARDLY
004   RDPORT-?   (M,DNA        )  001         ASSESS BY READING PORTFOLIOS:  MISSING, DOES NOT APPLY

CONDITIONING ID:              TRED0041
DESCRIPTION:                  HOW OFTEN DO YOU USE STUDENT SELF-REPORTS TO ASSESS STUDENTS IN READING?
GRADES/ASSESSMENTS:           N04, S04
GROUP LABEL:                  SELF_REP            LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                      T047008             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:             CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   SLFREP-1   (1            )  000         ASSESS BY STUDENT SELF-REPORTS:  ONCE OR TWICE A WEEK
002   SLFREP-2   (2            )  100         ASSESS BY STUDENT SELF-REPORTS:  ONCE OR TWICE A MONTH
003   SLFREP-3   (3,4          )  010         ASSESS BY STUDENT SELF-REP:  ONCE OR TWICE/YR, NEVER/HARDLY
004   SLFREP-?   (M,DNA        )  001         ASSESS BY STUDENT SELF-REPORTS:  MISSING, DOES NOT APPLY

CONDITIONING ID:              TRED0042
DESCRIPTION:                  HOW OFTEN DO YOU SEND OR TAKE THE CLASS TO THE LIBRARY?
GRADES/ASSESSMENTS:           N04, S04
GROUP LABEL:                  CLA2LIBR            LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                      T047101             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:             CLASS               NUMBER OF SPECIFICATION RECORDS:  5
001   CLALIB-1   (1            )  0000        TAKE CLASS TO LIBRARY:  ALMOST EVERY DAY
002   CLALIB-2   (2            )  1000        TAKE CLASS TO LIBRARY:  ONCE OR TWICE A WEEK
003   CLALIB-3   (3            )  0100        TAKE CLASS TO LIBRARY:  ONCE OR TWICE A MONTH
004   CLALIB-4   (4            )  0010        TAKE CLASS TO LIBRARY:  NEVER OR HARDLY EVER
005   CLALIB-5   (5,M,DNA      )  0001        TAKE CLASS TO LIBRARY:  NO LIBRARY, MISSING, DOES NOT APPLY

CONDITIONING ID:              TRED0043
DESCRIPTION:                  HOW OFTEN DO YOU ASSIGN STUDENTS TO READ A BOOK FROM THE LIBRARY?
GRADES/ASSESSMENTS:           N04, S04
GROUP LABEL:                  LIB_BOOK            LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                      T047102             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:             CLASS               NUMBER OF SPECIFICATION RECORDS:  5
001   CLALIB-1   (1            )  0000        ASSIGN STUDENTS TO READ LIBRARY BOOK:  ALMOST EVERY DAY
002   CLALIB-2   (2            )  1000        ASSIGN STUDENTS TO READ LIBRARY BOOK:  ONCE OR TWICE A WEEK
003   CLALIB-3   (3            )  0100        ASSIGN STUDENTS TO READ LIBRARY BOOK:  ONCE OR TWICE A MONTH
004   CLALIB-4   (4            )  0010        ASSIGN STUDENTS TO READ LIBRARY BOOK:  NEVER OR HARDLY EVER
005   CLALIB-5   (5,M,DNA      )  0001        ASSIGN STUDNTS TO RD LIB BOOK:  NO LIBRARY, MSSNG, NOT APPLY

CONDITIONING ID:              TRED0044
DESCRIPTION:                  ARE COMPUTERS AVAILABLE FOR USE BY STUDENTS IN READING CLASS?
GRADES/ASSESSMENTS:           N04, S04
GROUP LABEL:                  COMP4RDG            LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                      T047201             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:             CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   COMP-NA    (1            )  000         COMPUTERS IN READING CLASS:  NOT AVAILABLE
002   COMP-DIF   (2            )  100         COMPUTERS IN READING CLASS:  AVAILABLE BUT DIFF TO ACCESS
003   COMP-AVL   (3            )  010         COMPUTERS IN READING CLASS:  AVAILABLE IN THE CLASSROOM
004   COMP-?     (M,DNA        )  001         COMPUTERS IN READING CLASS:  MISSING, DOES NOT APPLY

CONDITIONING ID:              TRED0045
DESCRIPTION:                  OTHER TEACHING CERTIFICATION (DERIVED)
GRADES/ASSESSMENTS:           N04, S04
GROUP LABEL:                  CERTOTH             LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                      TRCERT              DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:             CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   MT/OTH-Y   (1            )  000         MIDDLE/JUNIOR HIGH/SECONDARY MATHEMATICS OR OTHER: YES
002   MATH-NO    (2            )  100         MIDDLE/JUNIOR HIGH/SECONDARY MATHEMATICS: NO
003   MATH-NS    (3            )  010         MIDDLE/JUNIOR HIGH/SECONDARY MATH: NOT OFFERED IN STATE
004   M/OTH-?    (M,DNA        )  001         MIDDLE/JR HIGH/SEC MATH: MISSING; OR OTHER: NO, NOT OFFERE

CONDITIONING ID:              TRED0046
DESCRIPTION:                  OTHER UNDERGRADUATE MAJOR (DERIVED)
GRADES/ASSESSMENTS:           N04, S04
```

Table F-2 (continued)
Conditioning Variables for Reading Main Samples

```
GROUP LABEL:                TRUMAJB          LENGTH OF CONTRAST FIELD      :  2
NAEP ID:                    TRUMAJB          DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS            NUMBER OF SPECIFICATION RECORDS:  3
001   UMAJO-Y    (1           )  00          OTHER OR (MATHEMATICS AND MATHEMATICS EDUCATION): YES
002   UMAJO-N    (2           )  10          OTHER AND (MATHEMATICS OR MATHEMATICS EDUCATION): MISSING
003   UMAJO-?    (M,DNA       )  01          OTHER AND (MATHEMATICS OR MATHEMATICS EDUCATION): MISSING

CONDITIONING ID:            TRED0047
DESCRIPTION:                OTHER GRADUATE MAJOR (DERIVED)
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                TRGMAJB          LENGTH OF CONTRAST FIELD      :  2
NAEP ID:                    TRGMAJB          DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS            NUMBER OF SPECIFICATION RECORDS:  3
001   GMAJO-Y    (1           )  00          OTHER OR (MATHEMATICS AND MATHEMATICS EDUCATION): YES
002   GMAJO-N    (2           )  10          OTHER AND (MATHEMATICS OR MATHEMATICS EDUCATION): MISSING
003   GMAJO-?    (M,DNA       )  01          OTHER AND (MATHEMATICS OR MATHEMATICS EDUCATION): MISSING

CONDITIONING ID:            TRED0048
DESCRIPTION:                SUM OF 'YES' RESPONSES TO TEACHER READING TRAINING VARIABLES (DERIVED)
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                TRTRAIN          LENGTH OF CONTRAST FIELD      :  4
NAEP ID:                    TRTRAIN          DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS            NUMBER OF SPECIFICATION RECORDS:  5
001   TRNS0-3    (1           )  0000        SUM(YES) = 0-3
002   TRNS4-6    (2           )  1000        SUM(YES) = 4-6
003   TRNS7      (3           )  0100        SUM(YES) = 7
004   TRNS8      (4           )  0010        SUM(YES) = 8
005   TRNS?      (M,DNA       )  0001        SUM(YES) = ?

CONDITIONING ID:            TRED0049
DESCRIPTION:                TEACHER EMPHASIS VARIABLE 1 (DERIVED)
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                TREMP1           LENGTH OF CONTRAST FIELD      :  2
NAEP ID:                    TREMP1           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS            NUMBER OF SPECIFICATION RECORDS:  3
001   TREMP1-1   (1           )  00          DECODING SKILLS AND PHONICS: BOTH HEAVY EMPHASIS
002   TREMP1-2   (2           )  10          DECODING SKILLS OR PHONICS: OTHERWISE
003   TREMP1-?   (DNA         )  01          DECODING SKILLS OR PHONICS: OTHERWISE

CONDITIONING ID:            TRED0050
DESCRIPTION:                TEACHER EMPHASIS VARIABLE 2 (DERIVED)
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                TREMP2           LENGTH OF CONTRAST FIELD      :  2
NAEP ID:                    TREMP2           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS            NUMBER OF SPECIFICATION RECORDS:  3
001   TREMP2-1   (1           )  00          VOCAB/INTEG OF RDG AND WRITING/WHOLE LANGUAGE: 2 OR MORE HEA
002   TREMP2-2   (2           )  10          VOCAB/INTEG OF READING AND WRITING/WHOLE LANGUAGE: OTHERWISE
003   TREMP2-?   (DNA         )  01          VOCAB/INTEG OF READING AND WRITING/WHOLE LANGUAGE: OTHERWISE

CONDITIONING ID:            TRED0051
DESCRIPTION:                TEACHER EMPHASIS VARIABLE 3 (DERIVED)
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                TREMP3           LENGTH OF CONTRAST FIELD      :  2
NAEP ID:                    TREMP3           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS            NUMBER OF SPECIFICATION RECORDS:  3
001   TREMP3-1   (1           )  00          RDG ACROSS CONTENT AREAS AND INDIVID RDG PROGRAMS: BOTH HEAV
002   TREMP3-2   (2           )  10          RDG ACROSS CONTENT AREAS AND INDIVID RDG PROGRAMS: OTHERWISE
003   TREMP3-?   (DNA         )  01          RDG ACROSS CONTENT AREAS AND INDIVID RDG PROGRAMS: OTHERWISE
```

Table F-3
Conditioning Variables for Mathematics Main Samples

```
CONDITIONING ID:          BACK0001
DESCRIPTION:              GRAND MEAN
GRADES/ASSESSMENTS:       N04, S04, N08, S08, N12
GROUP LABEL:              OVERALL              LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                  BKSER                DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         SCALE                NUMBER OF SPECIFICATION RECORDS:  1
001   OVERALL     (a           )  1            GRAND MEAN

CONDITIONING ID:          BACK0002
DESCRIPTION:              GENDER (DERIVED)
GRADES/ASSESSMENTS:       N04, S04, N08, S08, N12
GROUP LABEL:              GENDER               LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                  DSEX                 DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS                NUMBER OF SPECIFICATION RECORDS:  2
001   MALE        (1           )  0            GENDER:  MALE
002   FEMALE      (2           )  1            GENDER:  FEMALE

CONDITIONING ID:          BACK0003
DESCRIPTION:              ETHNICITY/RACE (DERIVED)
GRADES/ASSESSMENTS:       N04, S04, N08, S08, N12
GROUP LABEL:              ETHNICTY             LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                  DRACE                DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS                NUMBER OF SPECIFICATION RECORDS:  4
001   WHIT/AOM    (1,5,6,M     )  000          ETHNICITY:  WHITE, AMERICAN INDIAN, UNCLASSIFIED, MISSING
002   BLACK       (2           )  100          ETHNICITY:  BLACK
003   HISPANIC    (3           )  010          ETHNICITY:  HISPANIC
004   ASIAN       (4           )  001          ETHNICITY:  ASIAN AMERICAN

CONDITIONING ID:          BACK0004
DESCRIPTION:              SIZE AND TYPE OF COMMUNITY
GRADES/ASSESSMENTS:       N04, N08, N12
GROUP LABEL:              STOC                 LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                  STOC                 DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS                NUMBER OF SPECIFICATION RECORDS:  3
001   LO_METRO    (2           )  00           STOC:  LOW METROPOLITAN
002   HI_METRO    (3           )  10           STOC:  HIGH METROPOLITAN
003   STOC-OTH    (1,4-7,M     )  01           STOC:  OTHER

CONDITIONING ID:          BACK0006
DESCRIPTION:              REGION OF THE COUNTRY
GRADES/ASSESSMENTS:       N04, N08, N12
GROUP LABEL:              REGION               LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                  REGION               DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS                NUMBER OF SPECIFICATION RECORDS:  4
001   N_EAST      (1           )  000          REGION:  NORTHEAST
002   S_EAST      (2           )  100          REGION:  SOUTHEAST
003   CENTRAL     (3           )  010          REGION:  CENTRAL
004   WEST        (4           )  001          REGION:  WEST

CONDITIONING ID:          BACK0007
DESCRIPTION:              PARENTS' HIGHEST LEVEL OF EDUCATION
GRADES/ASSESSMENTS:       N04, S04, N08, S08, N12
GROUP LABEL:              PARED                LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                  PARED                DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS                NUMBER OF SPECIFICATION RECORDS:  5
001   <HI_SCH     (1           )  0000         PARED:  LESS THAN HIGH SCHOOL
002   HS_GRAD     (2           )  1000         PARED:  HIGH SCHOOL GRADUATE
003   POST_HS     (3           )  0100         PARED:  POST HIGH SCHOOL
004   COL_GRAD    (4           )  0010         PARED:  COLLEGE GRADUATE
005   PARED-?     (M,IDK       )  0001         PARED:  MISSING, I DON'T KNOW

CONDITIONING ID:          BACK0008
DESCRIPTION:              ITEMS IN THE HOME (NEWSPAPER, > 25 BOOKS, ENCYCLOPEDIA, MAGAZINES)
GRADES/ASSESSMENTS:       N04, S04, N08, S08, N12
```

BEST COPY AVAILABLE

Table F-3 (continued)
Conditioning Variables for Mathematics Main Samples

```
GROUP LABEL:              HOMEITMS          LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                  HOMEEN2           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  3
001   HITEM<=2   (1,M       )  00           ITEMS IN HOME:  ZERO TO TWO ITEMS, MISSING
002   HITEM=3    (2         )  10           ITEMS IN HOME:  THREE ITEMS
003   HITEM=4    (3         )  01           ITEMS IN HOME:  FOUR ITEMS

CONDITIONING ID:          BACK0009
DESCRIPTION:              HOURS OF TV WATCHING (LINEAR)
GRADES/ASSESSMENTS:       N04, S04, N08, S08, N12
GROUP LABEL:              TVWATCHL          LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                  B001801           DEGREES OF FREEDOM PER CONTRAST:  6
TYPE OF CONTRAST:         SCALE             NUMBER OF SPECIFICATION RECORDS:  7
001   TV-LIN1    (1         )  0            TV WATCHING (LINEAR):  NONE
002   TV-LIN2    (2         )  1            TV WATCHING (LINEAR):  ONE HOUR OR LESS PER DAY
003   TV-LIN3    (3         )  2            TV WATCHING (LINEAR):  TWO HOURS PER DAY
004   TV-LIN4    (4,M       )  3            TV WATCHING (LINEAR):  THREE HOURS PER DAY
005   TV-LIN5    (5         )  4            TV WATCHING (LINEAR):  FOUR HOURS PER DAY
006   TV-LIN6    (6         )  5            TV WATCHING (LINEAR):  FIVE HOURS PER DAY
007   TV-LIN7    (7         )  6            TV WATCHING (LINEAR):  SIX OR MORE HOURS PER DAY

CONDITIONING ID:          BACK0010
DESCRIPTION:              HOURS OF TV WATCHING (QUADRATIC)
GRADES/ASSESSMENTS:       N04, S04, N08, S08, N12
GROUP LABEL:              TVWATCHQ          LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                  B001801           DEGREES OF FREEDOM PER CONTRAST:  6
TYPE OF CONTRAST:         SCALE             NUMBER OF SPECIFICATION RECORDS:  7
001   TV-QUAD1   (1         )  00           TV WATCHING (QUADRATIC):  NONE
002   TV-QUAD2   (2         )  01           TV WATCHING (QUADRATIC):  ONE HOUR OR LESS PER DAY
003   TV-QUAD3   (3         )  04           TV WATCHING (QUADRATIC):  TWO HOURS PER DAY
004   TV-QUAD4   (4,M       )  09           TV WATCHING (QUADRATIC):  THREE HOURS PER DAY
005   TV-QUAD5   (5         )  16           TV WATCHING (QUADRATIC):  FOUR HOURS PER DAY
006   TV-QUAD6   (6         )  25           TV WATCHING (QUADRATIC):  FIVE HOURS PER DAY
007   TV-QUAD7   (7         )  36           TV WATCHING (QUADRATIC):  SIX OR MORE HOURS PER DAY

CONDITIONING ID:          BACK0011
DESCRIPTION:              HOME LANG MINORITY (HOW OFTEN PEOPLE IN HOME SPEAK LANG OTHER THAN ENGLISH?)
GRADES/ASSESSMENTS:       N04, S04, N08, S08, N12
GROUP LABEL:              HOMELANG          LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                  B003201           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  2
001   HL-NEV/?   (1,M       )  0            HOME LANGUAGE MINORITY:  NEVER, MISSING
002   HL-SM/AL   (2,3       )  1            HOME LANGUAGE MINORITY:  SOMTIMES, ALWAYS

CONDITIONING ID:          BACK0012
DESCRIPTION:              HOMEWORK ASSIGNED? (GRADE 4)
GRADES/ASSESSMENTS:       N04, S04
GROUP LABEL:              HW-CORE4          LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                  B006601           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  3
001   HW4-MISS   (M         )  00           HOMEWORK ASSIGNED?:  MISSING
002   HW4-NONE   (1         )  10           HOMEWORK ASSIGNED?:  NO HOMEWORK ASSIGNED
003   HW4-YES    (2-5       )  01           HOMEWORK ASSIGNED?:  YES

CONDITIONING ID:          BACK0013
DESCRIPTION:              AMOUNT OF HOMEWORK (LINEAR) (GRADE 4)
GRADES/ASSESSMENTS:       N04, S04
GROUP LABEL:              HMWRKL4           LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                  B006601           DEGREES OF FREEDOM PER CONTRAST:  3
TYPE OF CONTRAST:         SCALE             NUMBER OF SPECIFICATION RECORDS:  4
001   HW4-LIN1   (1,2,M     )  0            AMOUNT OF HOMEWORK (LINEAR):  DON'T HAVE, DON'T DO, MISSING
002   HW4-LIN2   (3         )  1            AMOUNT OF HOMEWORK (LINEAR):  ONE HALF HOUR
003   HW4-LIN3   (4         )  2            AMOUNT OF HOMEWORK (LINEAR):  ONE HOUR
004   HW4-LIN4   (5         )  3            AMOUNT OF HOMEWORK (LINEAR):  MORE THAN ONE HOUR
```

Table F-3 (continued)
Conditioning Variables for Mathematics Main Samples

```
CONDITIONING ID:           BACK0014
DESCRIPTION:               AMOUNT OF HOMEWORK (QUADRATIC) (GRADE 4)
GRADES/ASSESSMENTS:        N04, S04
GROUP LABEL:               HMWRKQ4          LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                   B006601          DEGREES OF FREEDOM PER CONTRAST:   3
TYPE OF CONTRAST:          SCALE            NUMBER OF SPECIFICATION RECORDS:   4
001   HW4QUAD1   (1,2,M      )  0           AMOUNT OF HOMEWORK (QUADRATIC):  DON'T HAVE, DON'T DO, MISS
002   HW4QUAD2   (3          )  1           AMOUNT OF HOMEWORK (QUADRATIC):  ONE HALF HOUR
003   HW4QUAD3   (4          )  4           AMOUNT OF HOMEWORK (QUADRATIC):  ONE HOUR
004   HW4QUAD4   (5          )  9           AMOUNT OF HOMEWORK (QUADRATIC):  MORE THAN ONE HOUR

CONDITIONING ID:           BACK0015
DESCRIPTION:               HOMEWORK ASSIGNED?
GRADES/ASSESSMENTS:        N08, S08, N12
GROUP LABEL:               HW-CORE          LENGTH OF CONTRAST FIELD       :   2
NAEP ID:                   B003901          DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS            NUMBER OF SPECIFICATION RECORDS:   3
001   HWC-MISS   (M          )  00          HOMEWORK ASSIGNED?:  MISSING
002   HWC-NONE   (1          )  10          HOMEWORK ASSIGNED?:  NO HOMEWORK ASSIGNED
003   HWC-YES    (2-6        )  01          HONEWORK ASSIGNED?:  YES

CONDITIONING ID:           BACK0016
DESCRIPTION:               AMOUNT OF HOMEWORK (LINEAR)
GRADES/ASSESSMENTS:        N08, S08, N12
GROUP LABEL:               HMWRKL           LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                   B003901          DEGREES OF FREEDOM PER CONTRAST:   4
TYPE OF CONTRAST:          SCALE            NUMBER OF SPECIFICATION RECORDS:   5
001   HW-LIN1    (1,2,M      )  0           AMOUNT OF HOMEWORK (LINEAR):  DON'T HAVE, DON'T DO, MISSING
002   HW-LIN2    (3          )  1           AMOUNT OF HOMEWORK (LINEAR):  ONE HALF HOUR
003   HW-LIN3    (4          )  2           AMOUNT OF HOMEWORK (LINEAR):  ONE HOUR
004   HW-LIN4    (5          )  3           AMOUNT OF HOMEWORK (LINEAR):  TWO HOURS
005   HW-LIN5    (6          )  4           AMOUNT OF HOMEWORK (LINEAR):  MORE THAN TWO HOURS

CONDITIONING ID:           BACK0017
DESCRIPTION:               AMOUNT OF HOMEWORK (QUADRATIC)
GRADES/ASSESSMENTS:        N08, S08, N12
GROUP LABEL:               HMWRKQ           LENGTH OF CONTRAST FIELD       :   2
NAEP ID:                   B003901          DEGREES OF FREEDOM PER CONTRAST:   4
TYPE OF CONTRAST:          SCALE            NUMBER OF SPECIFICATION RECORDS:   5
001   HW-QUAD1   (1,2,M      )  00          AMOUNT OF HOMEWORK (QUADRATIC):  DON'T HAVE, DON'T DO, MISS
002   HW-QUAD2   (3          )  01          AMOUNT OF HOMEWORK (QUADRATIC):  ONE HALF HOUR
003   HW-QUAD3   (4          )  04          AMOUNT OF HOMEWORK (QUADRATIC):  ONE HOUR
004   HW-QUAD4   (5          )  09          AMOUNT OF HOMEWORK (QUADRATIC):  TWO HOURS
005   HW-QUAD5   (6          )  16          AMOUNT OF HOMEWORK (QUADRATIC):  MORE THAN TWO HOURS

CONDITIONING ID:           BACK0018
DESCRIPTION:               PERCENT WHITE STUDENTS IN SCHOOL
GRADES/ASSESSMENTS:        N04, S04, N08, S08, N12
GROUP LABEL:               %WHITE           LENGTH OF CONTRAST FIELD       :   2
NAEP ID:                   PCTWHT           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS            NUMBER OF SPECIFICATION RECORDS:   3
001   PREDOM/?   (80-110,M   )  00          PREDOMINANTLY WHITE, MISSING
002   MINORITY   (0-49       )  10          WHITE MINORITY
003   INTEGRAT   (50-79      )  01          INTEGRATED

CONDITIONING ID:           BACK0019
DESCRIPTION:               MODAL AGE, MODAL GRADE (DERIVED)
GRADES/ASSESSMENTS:        N04, N08, N12
GROUP LABEL:               AGE/GRAD         LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                   MODGRAG          DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS            NUMBER OF SPECIFICATION RECORDS:   5
001   <MA/=MG    (1          )  0000        LESS THAN MODAL AGE, MODAL GRADE
002   =MA/<MG    (2          )  1000        MODAL AGE, LESS THAN MODAL GRADE
003   =MA/=MG    (3          )  0100        MODAL AGE, MODAL GRADE, MISSING
004   =MA/>MG    (4          )  0010        MODAL AGE, GREATER THAN MODAL GRADE
```

Table F-3 (continued)
Conditioning Variables for Mathematics Main Samples

```
005   >MA/=MG    (5          )  0001        GREATER THAN MODAL AGE, MODAL GRADE

CONDITIONING ID:            BACK0020
DESCRIPTION:                SCHOOL TYPE:  PUBLIC/NON-PUBLIC
GRADES/ASSESSMENTS:         N04, N08, N12
GROUP LABEL:                SCH_TYPE            LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                    SCHTYPE             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   2
001   PUBLIC     (1           )  0              PUBLIC SCHOOL
002   NON_PUBL   (2-5,M       )  1              PRIVATE, CATHOLIC, BIA, DEPT OF DEFENSE, MISS

CONDITIONING ID:            BACK0021
DESCRIPTION:                SINGLE/MULTIPLE PARENT(S) AT HOME
GRADES/ASSESSMENTS:         N04, S04, N08, S08, N12
GROUP LABEL:                PARENTS             LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                    SINGLEP             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   2
001   NOT2PARS   (2-4,M       )  0              NOT TWO PARENTS, MISSING
002   2PARENTS   (1           )  1              BOTH FATHER AND MOTHER AT HOME

CONDITIONING ID:            BACK0022
DESCRIPTION:                MOTHER AT HOME
GRADES/ASSESSMENTS:         N04, S04, N08, S08, N12
GROUP LABEL:                MOM@HOME            LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                    B005601             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   2
001   MOM@HM-N   (2,M         )  0              MOTHER AT HOME:  NO, MISSING
002   MOM@HM-Y   (1           )  1              MOTHER AT HOME:  YES

CONDITIONING ID:            BACK0023
DESCRIPTION:                PAGES READ FOR SCHOOL AND HOMEWORK EACH DAY (CONTRAST 1)
GRADES/ASSESSMENTS:         N04, S04, N08, S08, N12
GROUP LABEL:                PGSREAD1            LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                    B001101             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   2
001   <=5_PGS    (5,M         )  0              PAGES READ (1):  15 OR FEWER PAGES, MISSING
002   >=6_PGS    (1-4         )  1              PAGES READ (1):  MORE THAN 20, 16-20, 11-15, 6-10 PAGES

CONDITIONING ID:            BACK0024
DESCRIPTION:                PAGES READ FOR SCHOOL AND HOMEWORK EACH DAY (CONTRAST 2)
GRADES/ASSESSMENTS:         N04, S04, N08, S08, N12
GROUP LABEL:                PGSREAD2            LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                    B001101             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   2
001   <=10_PGS   (4,5,M       )  0              PAGES READ (2):  6-10 PAGES, 5 OR FEWER PAGES, MISSING
002   >=11_PGS   (1-3         )  1              PAGES READ (2):  MORE THAN 20, 16-20 PAGES, 11-15 PAGES

CONDITIONING ID:            BACK0025
DESCRIPTION:                WENT TO PRESCHOOL?
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                PRESCH              LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                    B004201             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   2
001   PRESCH-N   (2,3,IDK,M   )  0              WENT TO PRESCHOOL?:  NO, I DON'T KNOW, MISSING
002   PRESCH-Y   (1           )  1              WENT TO PRESCHOOL?:  YES

CONDITIONING ID:            BACK0026
DESCRIPTION:                DAYS OF SCHOOL MISSED LAST MONTH
GRADES/ASSESSMENTS:         N08, S08, N12
GROUP LABEL:                SCH_MISS            LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                    S004001             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   2
001   >=3_DAYS   (3-5,M       )  0              DAYS SCHOOL MISSED:  3 OR 4, 5 TO 10, MORE THAN 10, MISSING
002   <=2_DAYS   (1,2         )  1              DAYS SCHOOL MISSED:  NONE, 1 OR 2 DAYS
```

Table F-3 (continued)
Conditioning Variables for Mathematics Main Samples

```
CONDITIONING ID:            BACK0027
DESCRIPTION:                HIGH SCHOOL PROGRAM
GRADES/ASSESSMENTS:         N12
GROUP LABEL:                HS_PROG          LENGTH OF CONTRAST FIELD      :  2
NAEP ID:                    B005001          DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS            NUMBER OF SPECIFICATION RECORDS:  3
001   GENERL/?    (1,M        )  00          HS PROGRAM:  GENERAL, MISSING
002   COL_PREP    (2          )  10          HS PROGRAM:  COLLEGE PREPARATORY
003   VOC/TECH    (3          )  01          HS PROGRAM:  VOCATIONAL, TECHNICAL

CONDITIONING ID:            BACK0028
DESCRIPTION:                NUMBER OF SEMESTERS ENGLISH/LITERATURE/WRITING (MISSING/NOT-MISSING)
GRADES/ASSESSMENTS:         N12
GROUP LABEL:                #SEM_ENG         LENGTH OF CONTRAST FIELD      :  1
NAEP ID:                    B007101          DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS            NUMBER OF SPECIFICATION RECORDS:  2
001   #ENG-?      (M          )  0           NUMBER OF SEMESTERS ENGLISH MISSING
002   #ENG-Y      (1-9        )  1           NUMBER OF SEMESTERS ENGLISH NOT-MISSING

CONDITIONING ID:            BACK0029
DESCRIPTION:                NUMBER OF SEMESTERS ENGLISH/LITERATURE/WRITING (LINEAR)
GRADES/ASSESSMENTS:         N12
GROUP LABEL:                #ENG-LIN         LENGTH OF CONTRAST FIELD      :  1
NAEP ID:                    B007101          DEGREES OF FREEDOM PER CONTRAST:  8
TYPE OF CONTRAST:           SCALE            NUMBER OF SPECIFICATION RECORDS: 10
001   #ENG-L01    (1          )  1           NUMBER OF SEMESTERS ENGLISH LINEAR
002   #ENG-L02    (2          )  2           NUMBER OF SEMESTERS ENGLISH LINEAR
003   #ENG-L03    (3          )  3           NUMBER OF SEMESTERS ENGLISH LINEAR
004   #ENG-L04    (4          )  4           NUMBER OF SEMESTERS ENGLISH LINEAR
005   #ENG-L05    (5          )  5           NUMBER OF SEMESTERS ENGLISH LINEAR
006   #ENG-L06    (6          )  6           NUMBER OF SEMESTERS ENGLISH LINEAR
007   #ENG-L07    (7          )  7           NUMBER OF SEMESTERS ENGLISH LINEAR
008   #ENG-L08    (8          )  8           NUMBER OF SEMESTERS ENGLISH LINEAR
009   #ENG-L09    (9          )  9           NUMBER OF SEMESTERS ENGLISH LINEAR
010   #ENG-L10    (M          )  0           NUMBER OF SEMESTERS ENGLISH LINEAR (MISSING)

CONDITIONING ID:            BACK0030
DESCRIPTION:                NUMBER OF SEMESTERS MATHEMATICS (MISSING/NOT-MISSING)
GRADES/ASSESSMENTS:         N12
GROUP LABEL:                #SEM_MAT         LENGTH OF CONTRAST FIELD      :  1
NAEP ID:                    B007102          DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS            NUMBER OF SPECIFICATION RECORDS:  2
001   #MATH-?     (M          )  0           NUMBER OF SEMESTERS MATHEMATICS MISSING
002   #MATH-Y     (1-9        )  1           NUMBER OF SEMESTERS MATHEMATICS NOT-MISSING

CONDITIONING ID:            BACK0031
DESCRIPTION:                NUMBER OF SEMESTERS MATHEMATICS/LITERATURE/WRITING (LINEAR)
GRADES/ASSESSMENTS:         N12
GROUP LABEL:                #MAT-LIN         LENGTH OF CONTRAST FIELD      :  1
NAEP ID:                    B007102          DEGREES OF FREEDOM PER CONTRAST:  8
TYPE OF CONTRAST:           SCALE            NUMBER OF SPECIFICATION RECORDS: 10
001   #MAT-L01    (1          )  1           NUMBER OF SEMESTERS MATHEMATICS LINEAR
002   #MAT-L02    (2          )  2           NUMBER OF SEMESTERS MATHEMATICS LINEAR
003   #MAT-L03    (3          )  3           NUMBER OF SEMESTERS MATHEMATICS LINEAR
004   #MAT-L04    (4          )  4           NUMBER OF SEMESTERS MATHEMATICS LINEAR
005   #MAT-L05    (5          )  5           NUMBER OF SEMESTERS MATHEMATICS LINEAR
006   #MAT-L06    (6          )  6           NUMBER OF SEMESTERS MATHEMATICS LINEAR
007   #MAT-L07    (7          )  7           NUMBER OF SEMESTERS MATHEMATICS LINEAR
008   #MAT-L08    (8          )  8           NUMBER OF SEMESTERS MATHEMATICS LINEAR
009   #MAT-L09    (9          )  9           NUMBER OF SEMESTERS MATHEMATICS LINEAR
010   #MAT-L10    (M          )  0           NUMBER OF SEMESTERS MATHEMATICS LINEAR (MISSING)

CONDITIONING ID:            BACK0032
DESCRIPTION:                NUMBER OF SEMESTERS SCIENCE (MISSING/NOT-MISSING)
GRADES/ASSESSMENTS:         N12
```

Table F-3 (continued)
Conditioning Variables for Mathematics Main Samples

```
GROUP LABEL:                  #SEM_SCI          LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                      B007103           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:             CLASS             NUMBER OF SPECIFICATION RECORDS:   2
001   #SCI-?     (M             )  0            NUMBER OF SEMESTERS SCIENCE MISSING
002   #SCI-Y     (1-9           )  1            NUMBER OF SEMESTERS SCIENCE NOT-MISSING

CONDITIONING ID:              BACK0033
DESCRIPTION:                  NUMBER OF SEMESTERS SCIENCE/LITERATURE/WRITING (LINEAR)
GRADES/ASSESSMENTS:           N12
GROUP LABEL:                  #SCI-LIN          LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                      B007103           DEGREES OF FREEDOM PER CONTRAST:   8
TYPE OF CONTRAST:             SCALE             NUMBER OF SPECIFICATION RECORDS:  10
001   #SCI-L01   (1             )  1            NUMBER OF SEMESTERS SCIENCE LINEAR
002   #SCI-L02   (2             )  2            NUMBER OF SEMESTERS SCIENCE LINEAR
003   #SCI-L03   (3             )  3            NUMBER OF SEMESTERS SCIENCE LINEAR
004   #SCI-L04   (4             )  4            NUMBER OF SEMESTERS SCIENCE LINEAR
005   #SCI-L05   (5             )  5            NUMBER OF SEMESTERS SCIENCE LINEAR
006   #SCI-L06   (6             )  6            NUMBER OF SEMESTERS SCIENCE LINEAR
007   #SCI-L07   (7             )  7            NUMBER OF SEMESTERS SCIENCE LINEAR
008   #SCI-L08   (8             )  8            NUMBER OF SEMESTERS SCIENCE LINEAR
009   #SCI-L09   (9             )  9            NUMBER OF SEMESTERS SCIENCE LINEAR
010   #SCI-L10   (M             )  0            NUMBER OF SEMESTERS SCIENCE LINEAR (MISSING)

CONDITIONING ID:              BACK0034
DESCRIPTION:                  NUMBER OF SEMESTERS HISTORY/SOCIAL STUDIES/GEOGRAPHY (MISSING/NOT-MISSING)
GRADES/ASSESSMENTS:           N12
GROUP LABEL:                  #SEM_HIS          LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                      B007104           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:             CLASS             NUMBER OF SPECIFICATION RECORDS:   2
001   #HIS-?     (M             )  0            NUMBER OF SEMESTERS HISTORY MISSING
002   #HIS-Y     (1-9           )  1            NUMBER OF SEMESTERS HISTORY NOT-MISSING

CONDITIONING ID:              BACK0035
DESCRIPTION:                  NUMBER OF SEMESTERS HISTORY/SOCIAL STUDIES/GEOGRAPHY (LINEAR)
GRADES/ASSESSMENTS:           N12
GROUP LABEL:                  #HIS-LIN          LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                      B007104           DEGREES OF FREEDOM PER CONTRAST:   8
TYPE OF CONTRAST:             SCALE             NUMBER OF SPECIFICATION RECORDS:  10
001   #HIS-L01   (1             )  1            NUMBER OF SEMESTERS HISTORY LINEAR
002   #HIS-L02   (2             )  2            NUMBER OF SEMESTERS HISTORY LINEAR
003   #HIS-L03   (3             )  3            NUMBER OF SEMESTERS HISTORY LINEAR
004   #HIS-L04   (4             )  4            NUMBER OF SEMESTERS HISTORY LINEAR
005   #HIS-L05   (5             )  5            NUMBER OF SEMESTERS HISTORY LINEAR
006   #HIS-L06   (6             )  6            NUMBER OF SEMESTERS HISTORY LINEAR
007   #HIS-L07   (7             )  7            NUMBER OF SEMESTERS HISTORY LINEAR
008   #HIS-L08   (8             )  8            NUMBER OF SEMESTERS HISTORY LINEAR
009   #HIS-L09   (9             )  9            NUMBER OF SEMESTERS HISTORY LINEAR
010   #HIS-L10   (M             )  0            NUMBER OF SEMESTERS HISTORY LINEAR (MISSING)

CONDITIONING ID:              BACK0036
DESCRIPTION:                  NUMBER OF SEMESTERS FOREIGN LANGUAGE(S) (MISSING/NOT-MISSING)
GRADES/ASSESSMENTS:           N12
GROUP LABEL:                  #SEM_LAN          LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                      B007105           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:             CLASS             NUMBER OF SPECIFICATION RECORDS:   2
001   #FLANG-?   (M             )  0            NUMBER OF SEMESTERS FOREIGN LANGUAGE MISSING
002   #FLANG-Y   (1-9           )  1            NUMBER OF SEMESTERS FOREIGN LANGUAGE NOT-MISSING

CONDITIONING ID:              BACK0037
DESCRIPTION:                  NUMBER OF SEMESTERS FOREIGN LANGUAGE(S) (LINEAR)
GRADES/ASSESSMENTS:           N12
GROUP LABEL:                  #LAN-LIN          LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                      B007105           DEGREES OF FREEDOM PER CONTRAST:   8
TYPE OF CONTRAST:             SCALE             NUMBER OF SPECIFICATION RECORDS:  10
001   #LAN-L01   (1             )  1            NUMBER OF SEMESTERS FOREIGN LANGUAGE LINEAR
```

Table F-3 (continued)
Conditioning Variables for Mathematics Main Samples

```
002   #LAN-L02   (2          )  2          NUMBER OF SEMESTERS FOREIGN LANGUAGE LINEAR
003   #LAN-L03   (3          )  3          NUMBER OF SEMESTERS FOREIGN LANGUAGE LINEAR
004   #LAN-L04   (4          )  4          NUMBER OF SEMESTERS FOREIGN LANGUAGE LINEAR
005   #LAN-L05   (5          )  5          NUMBER OF SEMESTERS FOREIGN LANGUAGE LINEAR
006   #LAN-L06   (6          )  6          NUMBER OF SEMESTERS FOREIGN LANGUAGE LINEAR
007   #LAN-L07   (7          )  7          NUMBER OF SEMESTERS FOREIGN LANGUAGE LINEAR
008   #LAN-L08   (8          )  8          NUMBER OF SEMESTERS FOREIGN LANGUAGE LINEAR
009   #LAN-L09   (9          )  9          NUMBER OF SEMESTERS FOREIGN LANGUAGE LINEAR
010   #LAN-L10   (M          )  0          NUMBER OF SEMESTERS FOREIGN LANGUAGE LINEAR (MISSING)

CONDITIONING ID:          BACK0038
DESCRIPTION:              NUMBER OF SEMESTERS VOC/TECHNICAL/BUSINESS EDUCATION (MISSING/NOT-MISSING)
GRADES/ASSESSMENTS:       N12
GROUP LABEL:              #SEM_VOC             LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                  B007106              DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS                NUMBER OF SPECIFICATION RECORDS:  2
001   #VOC-?     (M          )  0          NUMBER OF SEMESTERS VOC/TECH MISSING
002   #VOC-Y     (1-9        )  1          NUMBER OF SEMESTERS VOC/TECH NOT-MISSING

CONDITIONING ID:          BACK0039
DESCRIPTION:              NUMBER OF SEMESTERS VOCATIONAL/TECHNICAL/BUSINESS EDUCATION (LINEAR)
GRADES/ASSESSMENTS:       N12
GROUP LABEL:              #VOC-LIN             LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                  B007106              DEGREES OF FREEDOM PER CONTRAST:  8
TYPE OF CONTRAST:         SCALE                NUMBER OF SPECIFICATION RECORDS: 10
001   #VOC-L01   (1          )  1          NUMBER OF SEMESTERS VOC/TECH LINEAR
002   #VOC-L02   (2          )  2          NUMBER OF SEMESTERS VOC/TECH LINEAR
003   #VOC-L03   (3          )  3          NUMBER OF SEMESTERS VOC/TECH LINEAR
004   #VOC-L04   (4          )  4          NUMBER OF SEMESTERS VOC/TECH LINEAR
005   #VOC-L05   (5          )  5          NUMBER OF SEMESTERS VOC/TECH LINEAR
006   #VOC-L06   (6          )  6          NUMBER OF SEMESTERS VOC/TECH LINEAR
007   #VOC-L07   (7          )  7          NUMBER OF SEMESTERS VOC/TECH LINEAR
008   #VOC-L08   (8          )  8          NUMBER OF SEMESTERS VOC/TECH LINEAR
009   #VOC-L09   (9          )  9          NUMBER OF SEMESTERS VOC/TECH LINEAR
010   #VOC-L10   (M          )  0          NUMBER OF SEMESTERS VOC/TECH LINEAR (MISSING)

CONDITIONING ID:          BACK0040
DESCRIPTION:              NUMBER OF SEMESTERS ART/MUSIC (MISSING/NOT-MISSING)
GRADES/ASSESSMENTS:       N12
GROUP LABEL:              #SEM_ART             LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                  B007107              DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS                NUMBER OF SPECIFICATION RECORDS:  2
001   #ART-?     (M          )  0          NUMBER OF SEMESTERS ART/MUSIC MISSING
002   #ART-Y     (1-9        )  1          NUMBER OF SEMESTERS ART/MUSIC NOT-MISSING

CONDITIONING ID:          BACK0041
DESCRIPTION:              NUMBER OF SEMESTERS ART/MUSIC (LINEAR)
GRADES/ASSESSMENTS:       N12
GROUP LABEL:              #ART-LIN             LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                  B007107              DEGREES OF FREEDOM PER CONTRAST:  8
TYPE OF CONTRAST:         SCALE                NUMBER OF SPECIFICATION RECORDS: 10
001   #ART-L01   (1          )  1          NUMBER OF SEMESTERS ART/MUSIC LINEAR
002   #ART-L02   (2          )  2          NUMBER OF SEMESTERS ART/MUSIC LINEAR
003   #ART-L03   (3          )  3          NUMBER OF SEMESTERS ART/MUSIC LINEAR
004   #ART-L04   (4          )  4          NUMBER OF SEMESTERS ART/MUSIC LINEAR
005   #ART-L05   (5          )  5          NUMBER OF SEMESTERS ART/MUSIC LINEAR
006   #ART-L06   (6          )  6          NUMBER OF SEMESTERS ART/MUSIC LINEAR
007   #ART-L07   (7          )  7          NUMBER OF SEMESTERS ART/MUSIC LINEAR
008   #ART-L08   (8          )  8          NUMBER OF SEMESTERS ART/MUSIC LINEAR
009   #ART-L09   (9          )  9          NUMBER OF SEMESTERS ART/MUSIC LINEAR
010   #ART-L10   (M          )  0          NUMBER OF SEMESTERS ART/MUSIC LINEAR (MISSING)

CONDITIONING ID:          BACK0042
DESCRIPTION:              BORN IN ONE OF THE 50 STATES
GRADES/ASSESSMENTS:       N04, S04, N08, S08, N12
```

Table F-3 (continued)
Conditioning Variables for Mathematics Main Samples

```
GROUP LABEL:              BORN_USA          LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                  B007801           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  2
001   USA-YES    (1         )  0            BORN IN THE USA:  YES
002   USA-NO/?   (2,M       )  1            BORN IN THE USA:  NO/MIS SING

CONDITIONING ID:          BACK0043
DESCRIPTION:              HOW MANY TIMES CHANGED SCHOOLS IN THE LAST TWO YEARS?
GRADES/ASSESSMENTS:       N04, S04, N08, S08
GROUP LABEL:              SCH_CHGS          LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                  B007301           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  4
001   CHGSCH=0   (1         )  000          CHANGED SCHOOLS (NONE)
002   CHGSCH=1   (2         )  100          CHANGED SCHOOLS ONCE
003   CHGSCH=2   (3         )  010          CHANGED SCHOOLS TWICE
004   CHGSCH3+   (4,M       )  001          CHANGED SCHOOLS 3 OR MORE TIMES, MISSING

CONDITIONING ID:          BACK0044
DESCRIPTION:              HOW MANY GRADES HAVE YOU GONE TO SCHOOL IN THIS STATE? (K-4)
GRADES/ASSESSMENTS:       N04, S04
GROUP LABEL:              GRDS_ST4          LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                  B007601           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  3
001   ST4GRD<1   (1,M       )  00           LESS THAN ONE GRADE IN THIS STATE, MISSING (K-4)
002   ST4GRD12   (2         )  10           ONE TO TWO GRADES IN THIS STATE (K-4)
003   ST4GRD3+   (3         )  01           THREE OR MORE GRADES IN THIS STATE (K-4)

CONDITIONING ID:          BACK0045
DESCRIPTION:              HOW OFTEN DO YOU DISCUSS THINGS STUDIED IN SCHOOL WITH SOMEONE AT HOME?
GRADES/ASSESSMENTS:       N04, S04, N08, S08, N12
GROUP LABEL:              DISC@HOM          LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                  B007401           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  4
001   DIS@HOM1   (1         )  000          DISCUSS AT HOME (ALMOST EVERYDAY)
002   DIS@HOM2   (2         )  100          DISCUSS AT HOME (ONCE OR TWICE A WEEK)
003   DIS@HOM3   (3         )  010          DISCUSS AT HOME (ONCE OR TWICE A MONTH)
004   DIS@HOM4   (4,M       )  001          DISCUSS AT HOME (NEVER OR HARDLY EVER, MISSING)

CONDITIONING ID:          BACK0046
DESCRIPTION:              HOW OFTEN DO USE A COMPUTER FOR SCHOOLWORK?
GRADES/ASSESSMENTS:       N04, S04, N08, S08, N12
GROUP LABEL:              COMP4SCH          LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                  B007501           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  5
001   CMP4SCH1   (1         )  0000         COMPUTER FOR SCHOOLWORK (ALMOST EVERYDAY)
002   CMP4SCH2   (2         )  1000         COMPUTER FOR SCHOOLWORK (ONCE OR TWICE A WEEK)
003   CMP4SCH3   (3         )  0100         COMPUTER FOR SCHOOLWORK (ONCE OR TWICE A MONTH)
004   CMP4SCH4   (4         )  0010         COMPUTER FOR SCHOOLWORK (NEVER OR HARDLY EVER)
005   CMP4SCH5   (M         )  0001         COMPUTER FOR SCHOOLWORK (MISSING)

CONDITIONING ID:          BACK0047
DESCRIPTION:              HOW MANY GRADES HAVE YOU GONE TO SCHOOL IN THIS STATE? (K-8)
GRADES/ASSESSMENTS:       N08, S08
GROUP LABEL:              GRDS_ST8          LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                  B007701           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  4
001   ST8GRD<1   (1,M       )  000          LESS THAN ONE GRADE IN THIS STATE, MISSING (K-8)
002   ST8GRD12   (2         )  100          ONE TO TWO GRADES IN THIS STATE (K-8)
003   ST8GRD35   (3         )  010          THREE TO FIVE GRADES IN THIS STATE (K-8)
004   ST8GRD>5   (4         )  001          MORE THAN FIVE GRADES IN THIS STATE (K-8)

CONDITIONING ID:          BACK0048
DESCRIPTION:              WHAT DO YOU EXPECT YOUR MAIN ACTIVITY WILL BE THE YEAR AFTER SCHOOL?
GRADES/ASSESSMENTS:       N12
GROUP LABEL:              MAINACT           LENGTH OF CONTRAST FIELD       :  5
```

Table F-3 (continued)
Conditioning Variables for Mathematics Main Samples

```
NAEP ID:                     B007201            DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:            CLASS              NUMBER OF SPECIFICATION RECORDS:   6
001   WORK_F/T   (1             )   00000       MAIN ACTIVITY:  WORK FULL-TIME
002   VOC/BUSI   (2             )   10000       MAIN ACTIVITY:  ATTEND VOC, TECH, BUSINESS SCHOOL
003   2-YR_COL   (3             )   01000       MAIN ACTIVITY:  2-YEAR COLLEGE
004   4-YR_COL   (4             )   00100       MAIN ACTIVITY:  4-YEAR COLLEGE
005   MILITARY   (5             )   00010       MAIN ACTIVITY:  MILITARY
006   OTHERACT   (6,M           )   00001       MAIN ACTIVITY:  OTHER, MISSING

CONDITIONING ID:             MATH0001
DESCRIPTION:                 SCHOOL LEVEL AVERAGE MATHEMATICS PROFICIENCY
GRADES/ASSESSMENTS:          N04, S04, N08, S08, N12
GROUP LABEL:                 SLP_MATH           LENGTH OF CONTRAST FIELD        :   1
NAEP ID:                     SCHMATH            DEGREES OF FREEDOM PER CONTRAST: 999
TYPE OF CONTRAST:            CLASS              NUMBER OF SPECIFICATION RECORDS:   2
001   SLP_MA-Y   (@             )   1           SCHOOL LEVEL AVERAGE MATHEMATICS PROFICIENCY NOT-MISSING
002   SLP_MA-?   (M             )   0           SCHOOL LEVEL AVERAGE MATHEMATICS PROFICIENCY MISSING

CONDITIONING ID:             MATH0002
DESCRIPTION:                 SCHOOL LEVEL AVERAGE MATHEMATICS PROFICIENCY
GRADES/ASSESSMENTS:          N04, S04, N08, S08, N12
GROUP LABEL:                 SLP_MATL           LENGTH OF CONTRAST FIELD        :   8
NAEP ID:                     SCHMATH            DEGREES OF FREEDOM PER CONTRAST: 999
TYPE OF CONTRAST:            SCALE              NUMBER OF SPECIFICATION RECORDS:   2
001   SLP_MA-L   (#             )   (F8.4)      SCHOOL LEVEL AVERAGE MATHEMATICS PROFICIENCY MEAN
002   SLP_MA-?   (M             )   0           SCHOOL LEVEL AVERAGE MATHEMATICS PROFICIENCY MISSING

CONDITIONING ID:             MATH0003
DESCRIPTION:                 HOW OFTEN DO MATH PROBLEMS FROM TEXTBOOKS (STUDENT)?
GRADES/ASSESSMENTS:          N04, S04, N08, S08, N12
GROUP LABEL:                 S_TXTBKS           LENGTH OF CONTRAST FIELD        :   4
NAEP ID:                     M811601            DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:            CLASS              NUMBER OF SPECIFICATION RECORDS:   5
001   S_TXTBK1   (1             )   0000        MATH FROM TEXTBOOKS (STUDENT):  ALMOST EVERY DAY
002   S_TXTBK2   (2             )   1000        MATH FROM TEXTBOOKS (STUDENT):  ONCE OR TWICE A WEEK
003   S_TXTBK3   (3             )   0100        MATH FROM TEXTBOOKS (STUDENT):  ONCE OR TWICE A MONTH
004   S_TXTBK4   (4             )   0010        MATH FROM TEXTBOOKS (STUDENT):  NEVER OR HARDLY EVER
005   S_TXTBK?   (M             )   0001        MATH FROM TEXTBOOKS (STUDENT):  MISSING

CONDITIONING ID:             MATH0004
DESCRIPTION:                 HOW OFTEN DO MATH PROBLEMS ON WORKSHEETS?
GRADES/ASSESSMENTS:          N04, S04, N08, S08, N12
GROUP LABEL:                 S_WRKSHS           LENGTH OF CONTRAST FIELD        :   4
NAEP ID:                     M811602            DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:            CLASS              NUMBER OF SPECIFICATION RECORDS:   5
001   S_WRKSH1   (1             )   0000        MATH FROM WORKSHEETS (STUDENT): ALMOST EVERY DAY
002   S_WRKSH2   (2             )   1000        MATH FROM WORKSHEETS (STUDENT): ONCE OR TWICE A WEEK
003   S_WRKSH3   (3             )   0100        MATH FROM WORKSHEETS (STUDENT): ONCE OR TWICE A MONTH
004   S_WRKSH4   (4             )   0010        MATH FROM WORKSHEETS (STUDENT): NEVER OR HARDLY EVER
005   S_WRKSH?   (M             )   0001        MATH FROM WORKSHEETS (STUDENT): MISSING

CONDITIONING ID:             MATH0005
DESCRIPTION:                 HOW OFTEN SOLVE MATH PROBLEMS IN SMALL GROUPS?
GRADES/ASSESSMENTS:          N04, S04, N08, S08, N12
GROUP LABEL:                 S_SMGRPS           LENGTH OF CONTRAST FIELD        :   4
NAEP ID:                     M811603            DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:            CLASS              NUMBER OF SPECIFICATION RECORDS:   5
001   S_SMGRP1   (1             )   0000        MATH IN SMALL GROUPS (STUDENT): ALMOST EVERY DAY
002   S_SMGRP2   (2             )   1000        MATH IN SMALL GROUPS (STUDENT): ONCE OR TWICE A WEEK
003   S_SMGRP3   (3             )   0100        MATH IN SMALL GROUPS (STUDENT): ONCE OR TWICE A MONTH
004   S_SMGRP4   (4             )   0010        MATH IN SMALL GROUPS (STUDENT): NEVER OR HARDLY EVER
005   S_SMGRP?   (M             )   0001        MATH IN SMALL GROUPS (STUDENT): MISSING

CONDITIONING ID:             MATH0006
DESCRIPTION:                 HOW OFTEN WORK WITH OBJECTS LIKE RULERS, BLOCKS, SHAPES? (STUDENT)
```

Table F-3 (continued)
Conditioning Variables for Mathematics Main Samples

```
GRADES/ASSESSMENTS:        N04, S04
GROUP LABEL:               S_OBJCTS          LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                   M811604           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:  5
001   S_OBJCT1   (1          )  0000         WORK WITH OBJECTS (STUDENT): ALMOST EVERY DAY
002   S_OBJCT2   (2          )  1000         WORK WITH OBJECTS (STUDENT): ONCE OR TWICE A WEEK
003   S_OBJCT3   (3          )  0100         WORK WITH OBJECTS (STUDENT): ONCE OR TWICE A MONTH
004   S_OBJCT4   (4          )  0010         WORK WITH OBJECTS (STUDENT): NEVER OR HARDLY EVER
005   S_OBJCT?   (M          )  0001         WORK WITH OBJECTS (STUDENT): MISSING

CONDITIONING ID:           MATH0007
DESCRIPTION:               HOW OFTEN WORK WITH MEASUREMENT INSTRUMENTS/GEOMETRIC SOLIDS? (STUDENT)
GRADES/ASSESSMENTS:        N08, S08  N12
GROUP LABEL:               S_MI&GS           LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                   M811608           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:  5
001   S_MI&GS1   (1          )  0000         MEASURE INSTRUMENTS/GEOM SOLIDS (STUDENT): ALMOST EVERY DAY
002   S_MI&GS2   (2          )  1000         MEASURE INSTRUMENTS/GEOM SOLIDS (STUDENT): ONCE OR TWICE/WK
003   S_MI&GS3   (3          )  0100         MEASURE INSTRUMENTS/GEOM SOLIDS (STUDENT): ONCE/TWICE/MNTH
004   S_MI&GS4   (4          )  0010         MEASURE INSTRUMENTS/GEOM SOLIDS (STUDENT): NEVER OR HARDLY
005   S_MI&GS?   (M          )  0001         MEASURE INSTRUMENTS/GEOM SOLIDS (STUDENT): MISSING

CONDITIONING ID:           MATH0008
DESCRIPTION:               HOW OFTEN USE A CALCULATOR (STUDENT)?
GRADES/ASSESSMENTS:        N04, S04, N08, S08, N12
GROUP LABEL:               S_CALCTR          LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                   M811605           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:  5
001   S_CALCR1   (1          )  0000         USE CALCULATOR (STUDENT): ALMOST EVERY DAY
002   S_CALCR2   (2          )  1000         USE CALCULATOR (STUDENT): ONCE OR TWICE A WEEK
003   S_CALCR3   (3          )  0100         USE CALCULATOR (STUDENT): ONCE OR TWICE A MONTH
004   S_CALCR4   (4          )  0010         USE CALCULATOR (STUDENT): NEVER OR HARDLY EVER
005   S_CALCR?   (M          )  0001         USE CALCULATOR (STUDENT): MISSING

CONDITIONING ID:           MATH0009
DESCRIPTION:               HOW OFTEN USE A COMPUTER (STUDENT)?
GRADES/ASSESSMENTS:        N04, S04, N08, S08, N12
GROUP LABEL:               S_COMPTR          LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                   M811606           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:  5
001   S_CMPTR1   (1          )  0000         USE COMPUTER (STUDENT): ALMOST EVERY DAY
002   S_CMPTR2   (2          )  1000         USE COMPUTER (STUDENT): ONCE OR TWICE A WEEK
003   S_CMPTR3   (3          )  0100         USE COMPUTER (STUDENT): ONCE OR TWICE A MONTH
004   S_CMPTR4   (4          )  0010         USE COMPUTER (STUDENT): NEVER OR HARDLY EVER
005   S_CMPTR?   (M          )  0001         USE COMPUTER (STUDENT): MISSING

CONDITIONING ID:           MATH0010
DESCRIPTION:               HOW OFTEN WRITE ABOUT SOLVING MATH PROBLEM (STUDENT)?
GRADES/ASSESSMENTS:        N08, S08, N12
GROUP LABEL:               S_PRBSOL          LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                   M811609           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:  5
001   S_PRBSL1   (1          )  0000         PROBLEM SOLVING (STUDENT): ALMOST EVERY DAY
002   S_PRBSL2   (2          )  1000         PROBLEM SOLVING (STUDENT): ONCE OR TWICE A WEEK
003   S_PRBSL3   (3          )  0100         PROBLEM SOLVING (STUDENT): ONCE OR TWICE A MONTH
004   S_PRBSL4   (4          )  0010         PROBLEM SOLVING (STUDENT): NEVER OR HARDLY EVER
005   S_PRBSL?   (M          )  0001         PROBLEM SOLVING (STUDENT): MISSING

CONDITIONING ID:           MATH0011
DESCRIPTION:               HOW OFTEN MAKE UP MATH PROBLEMS FOR OTHERS TO SOLVE? (STUDENT)
GRADES/ASSESSMENTS:        N08, S08, N12
GROUP LABEL:               S_MUPROB          LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                   M811610           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:  5
001   S_MUPRB1   (1          )  0000         STUDENTS MAKE UP PROBLEMS (STUDENT): ALMOST EVERY DAY
```

Table F-3 (continued)
Conditioning Variables for Mathematics Main Samples

```
002   S_MUPRB2   (2          )  1000        STUDENTS MAKE UP PROBLEMS (STUDENT): ONCE OR TWICE A WEEK
003   S_MUPRB3   (3          )  0100        STUDENTS MAKE UP PROBLEMS (STUDENT): ONCE OR TWICE A MONTH
004   S_MUPRB4   (4          )  0010        STUDENTS MAKE UP PROBLEMS (STUDENT): NEVER OR HARDLY EVER
005   S_MUPRB?   (M          )  0001        STUDENTS MAKE UP PROBLEMS (STUDENT): MISSING

CONDITIONING ID:            MATH0012
DESCRIPTION:                HOW OFTEN TAKE MATH TESTS? (STUDENT)
GRADES/ASSESSMENTS:         N04, S04, N08, S08, N12
GROUP LABEL:                S_MATTST            LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                    M811607             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   5
001   S_MATST1   (1          )  0000        TAKE MATH TESTS (STUDENT): ALMOST EVERY DAY
002   S_MATST2   (2          )  1000        TAKE MATH TESTS (STUDENT): ONCE OR TWICE A WEEK
003   S_MATST3   (3          )  0100        TAKE MATH TESTS (STUDENT): ONCE OR TWICE A MONTH
004   S_MATST4   (4          )  0010        TAKE MATH TESTS (STUDENT): NEVER OR HARDLY EVER
005   S_MATST?   (M          )  0001        TAKE MATH TESTS (STUDENT): MISSING

CONDITIONING ID:            MATH0013
DESCRIPTION:                HOW OFTEN WRITE REPORTS/DO PROJECTS? (STUDENT)
GRADES/ASSESSMENTS:         N08, S08, N12
GROUP LABEL:                S_REPPRJ            LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                    M811611             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   5
001   S_REPPJ1   (1          )  0000        REPORTS/PROJECTS (STUDENT): ALMOST EVERY DAY
002   S_REPPJ2   (2          )  1000        REPORTS/PROJECTS (STUDENT): ONCE OR TWICE A WEEK
003   S_REPPJ3   (3          )  0100        REPORTS/PROJECTS (STUDENT): ONCE OR TWICE A MONTH
004   S_REPPJ4   (4          )  0010        REPORTS/PROJECTS (STUDENT): NEVER OR HARDLY EVER
005   S_REPPJ?   (M          )  0001        REPORTS/PROJECTS (STUDENT): MISSING

CONDITIONING ID:            MATH0014
DESCRIPTION:                HAVE A CALCULATOR TO DO MATH SCHOOLWORK?
GRADES/ASSESSMENTS:         N04, S04, N08, S08, N12
GROUP LABEL:                HAVECALC            LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                    M811201             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   2
001   HVCALC-Y   (1          )  0           HAVE CALCULATOR (STUDENT):  YES
002   HVCALC-N   (2,M        )  1           HAVE CALCULATOR (STUDENT):  NO, MISSING

CONDITIONING ID:            MATH0015
DESCRIPTION:                HAVE YOU EVER USED A SCIENTIFIC CALCULATOR?
GRADES/ASSESSMENTS:         N08, S08, N12
GROUP LABEL:                SCI_CALC            LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                    M810401             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   2
001   SCICAL-Y   (1          )  0           USED SCIENTIFIC CALCULATOR:  YES
002   SCICAL-N   (2,M        )  1           USED SCIENTIFIC CALCULATOR:  NO, MISSING

CONDITIONING ID:            MATH0016
DESCRIPTION:                WHAT KIND OF MATH CLASS ARE YOU TAKING THIS YEAR? (GRADE 8)
GRADES/ASSESSMENTS:         N08, S08
GROUP LABEL:                MCLASS8             LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                    M810501             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   5
001   NO_MATH8   (1,M        )  0000       8TH GRADE MATH CLASS: NO MATH THIS YEAR, MISSING
002   8GR_MAT8   (2          )  1000       8TH GRADE MATH CLASS: 8TH GRADE MATH
003   PRE-ALG8   (3          )  0100       8TH GRADE MATH CLASS: PRE-ALGEBRA
004   ALGEBRA8   (4          )  0010       8TH GRADE MATH CLASS: ALGEBRA
005   OTHER_M8   (5          )  0001       8TH GRADE MATH CLASS: OTHER

CONDITIONING ID:            MATH0017
DESCRIPTION:                WHAT MATH CLASS WILL YOU TAKE IN 9TH GRADE? (GRADE 8)
GRADES/ASSESSMENTS:         N08, S08
GROUP LABEL:                MCLASS9             LENGTH OF CONTRAST FIELD       :   6
NAEP ID:                    M811701             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   7
```

Table F-3 (continued)
Conditioning Variables for Mathematics Main Samples

```
001   NO_MATH9   (1          )  000000       9TH GRADE MATH CLASS: WON'T TAKE MATH IN 9TH GRADE
002   BASIC9     (2          )  100000       9TH GRADE MATH CLASS: BASIC, GEN, BUSINESS, CONSUMER MATH
003   PRE-ALG9   (3          )  010000       9TH GRADE MATH CLASS: PRE-ALGEBRA
004   ALGEBRA9   (4          )  001000       9TH GRADE MATH CLASS: ALGEBRA
005   GEOMTRY9   (5          )  000100       9TH GRADE MATH CLASS: GEOMETRY
006   OTHER_M9   (6          )  000010       9TH GRADE MATH CLASS: OTHER
007   MCLASS9?   (7,M        )  000001       9TH GRADE MATH CLASS: MISSING

CONDITIONING ID:                 MATH0018
DESCRIPTION:                     ARE YOU TAKING MATH CLASSES THIS YEAR? (GRADE 12)
GRADES/ASSESSMENTS:              N12
GROUP LABEL:                     MCLASS12          LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                         M810901           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:                CLASS             NUMBER OF SPECIFICATION RECORDS:  2
001   MATH12-Y   (1          )  0            TAKING MATH? (GRADE 12): YES
002   MATH12-N   (2,M        )  1            TAKING MATH? (GRADE 12): NO, MISSING

CONDITIONING ID:                 MATH0019
DESCRIPTION:                     IN WHAT GRADE DID YOU TAKE FIRST-YEAR ALGEBRA?
GRADES/ASSESSMENTS:              N12
GROUP LABEL:                     ALG1_GRD          LENGTH OF CONTRAST FIELD       :  5
NAEP ID:                         M811801           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:                CLASS             NUMBER OF SPECIFICATION RECORDS:  6
001   ALG1-NO    (1          )  00000        GRADE FIRST-YEAR ALGEBRA TAKEN:  HAVEN'T TAKEN
002   ALG1-<9    (2          )  10000        GRADE FIRST-YEAR ALGEBRA TAKEN:  BEFORE 9TH GRADE
003   ALG1-9     (3          )  01000        GRADE FIRST-YEAR ALGEBRA TAKEN:  9TH GRADE
004   ALG1-10    (4          )  00100        GRADE FIRST-YEAR ALGEBRA TAKEN:  10TH GRADE
005   ALG1->10   (5          )  00010        GRADE FIRST-YEAR ALGEBRA TAKEN:  11TH OR 12TH GRADE
006   ALG1-?     (M          )  00001        GRADE FIRST-YEAR ALGEBRA TAKEN:  MISSING

CONDITIONING ID:                 MATH0020
DESCRIPTION:                     HOW MUCH TIME SPENT ON MATH HOMEWORK EACH DAY? (STUDENT)
GRADES/ASSESSMENTS:              N04, S04, N08, S08, N12
GROUP LABEL:                     S_MATHHW          LENGTH OF CONTRAST FIELD       :  7
NAEP ID:                         M811301           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:                CLASS             NUMBER OF SPECIFICATION RECORDS:  8
001   S_MATHW1   (1          )  0000000      AMOUNT MATH HOMEWORK (STUDENT):  NOT TAKING MATH
002   S_MATHW1   (2          )  1000000      AMOUNT MATH HOMEWORK (STUDENT):  NONE
003   S_MATHW2   (3          )  0100000      AMOUNT MATH HOMEWORK (STUDENT):  15 MINUTES
004   S_MATHW3   (4          )  0010000      AMOUNT MATH HOMEWORK (STUDENT):  30 MINUTES
005   S_MATHW4   (5          )  0001000      AMOUNT MATH HOMEWORK (STUDENT):  45 MINUTES
006   S_MATHW5   (6          )  0000100      AMOUNT MATH HOMEWORK (STUDENT):  1 HOUR
007   S_MATHW6   (7          )  0000010      AMOUNT MATH HOMEWORK (STUDENT):  MORE THAN 1 HOUR
008   S_MATHW?   (M          )  0000001      AMOUNT MATH HOMEWORK (STUDENT):  MISSING, DOES NOT APPLY

CONDITIONING ID:                 MATH0021
DESCRIPTION:                     DO YOU GET HELP IN MATH FROM SPECIAL TEACHERS?
GRADES/ASSESSMENTS:              N04, S04
GROUP LABEL:                     MATHHELP          LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                         M811401           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:                CLASS             NUMBER OF SPECIFICATION RECORDS:  2
001   MATHLP-Y   (1          )  0            HELP IN MATH:  YES
002   MATHLP-N   (2,M        )  1            HELP IN MATH:  NO, MISSING

CONDITIONING ID:                 MATH0022
DESCRIPTION:                     AGREE/DISAGREE: I LIKE MATH (GRADE 4)
GRADES/ASSESSMENTS:              N04, S04
GROUP LABEL:                     LIKEMAT4          LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                         M811101           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:                CLASS             NUMBER OF SPECIFICATION RECORDS:  3
001   LKMATH-A   (1          )  00           I LIKE MATH (GRADE 4):  AGREE
002   LKMATH-U   (2          )  10           I LIKE MATH (GRADE 4):  UNDECIDED
003   LKMATH-D   (3,M        )  01           I LIKE MATH (GRADE 4):  DISAGREE, MISSING

CONDITIONING ID:                 MATH0023
```

Table F-3 (continued)
Conditioning Variables for Mathematics Main Samples

```
DESCRIPTION:              AGREE/DISAGREE: I AM GOOD AT MATH (GRADE 4)
GRADES/ASSESSMENTS:       N04, S04
GROUP LABEL:              GOODMAT4          LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                  M811103           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  3
001   GDMATH-A   (1         )  00           I AM GOOD AT MATH (GRADE 4): AGREE
002   GDMATH-U   (2         )  10           I AM GOOD AT MATH (GRADE 4): UNDECIDED
003   GDMATH-D   (3,M       )  01           I AM GOOD AT MATH (GRADE 4): DISAGREE, MISSING

CONDITIONING ID:          MATH0024
DESCRIPTION:              AGREE/DISAGREE: UNDERSTAND MOST OF MATH CLASS (GRADE 4)
GRADES/ASSESSMENTS:       N04, S04
GROUP LABEL:              USTDMAT4          LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                  M811106           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  3
001   USMATH-A   (1         )  00           UNDERSTAND MATH (GRADE 4): AGREE
002   USMATH-U   (2         )  10           UNDERSTAND MATH (GRADE 4): UNDECIDED
003   USMATH-D   (3,M       )  01           UNDERSTAND MATH (GRADE 4): DISAGREE, MISSING

CONDITIONING ID:          MATH0025
DESCRIPTION:              AGREE/DISAGREE: MATH IS MORE FOR BOYS THAN FOR GIRLS (GRADE 4)
GRADES/ASSESSMENTS:       N04, S04
GROUP LABEL:              MAT4BOY4          LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                  M811104           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  3
001   MA4BOY-A   (1,M       )  00           MATH FOR BOYS (GRADE 4): AGREE, MISSING
002   MA4BOY-U   (2         )  10           MATH FOR BOYS (GRADE 4): UNDECIDED
003   MA4BOY-D   (3         )  01           MATH FOR BOYS (GRADE 4): DISAGREE

CONDITIONING ID:          MATH0026
DESCRIPTION:              AGREE/DISAGREE: MATH MOSTLY MEMORIZING FACTS (GRADE 4)
GRADES/ASSESSMENTS:       N04, S04
GROUP LABEL:              MATMEMF4          LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                  M811107           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  3
001   MATMEM-A   (1         )  00           MATH IS MEMORIZING FACTS (GRADE 4): AGREE
002   MATMEM-U   (2         )  10           MATH IS MEMORIZING FACTS (GRADE 4): UNDECIDED
003   MATMEM-D   (3,M       )  01           MATH IS MEMORIZING FACTS (GRADE 4): DISAGREE, MISSING

CONDITIONING ID:          MATH0027
DESCRIPTION:              AGREE/DISAGREE: PEOPLE USE MATH IN JOBS (GRADE 4)
GRADES/ASSESSMENTS:       N04, S04
GROUP LABEL:              MATJOBS4          LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                  M811102           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  3
001   MATJOB-A   (1         )  00           USE MATH IN JOBS (GRADE 4): AGREE
002   MATJOB-U   (2         )  10           USE MATH IN JOBS (GRADE 4): UNDECIDED
003   MATJOB-D   (3,M       )  01           USE MATH IN JOBS (GRADE 4): DISAGREE, MISSING

CONDITIONING ID:          MATH0028
DESCRIPTION:              AGREE/DISAGREE: MATH USED FOR SOLVING EVERYDAY PROBLEMS (GRADE 4)
GRADES/ASSESSMENTS:       N04, S04
GROUP LABEL:              MAT4PRB4          LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                  M811105           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  3
001   MA4PRB-A   (1         )  00           USE MATH FOR SOLVING PROBLEMS (GRADE 4): AGREE
002   MA4PRB-U   (2         )  10           USE MATH FOR SOLVING PROBLEMS (GRADE 4): UNDECIDED
003   MA4PRB-D   (3,M       )  01           USE MATH FOR SOLVING PROBLEMS (GRADE 4): DISAGREE, MISSING

CONDITIONING ID:          MATH0029
DESCRIPTION:              AGREE/DISAGREE: IF CHOICE, WOULD NOT STUDY MORE MATH (GRADE 4)
GRADES/ASSESSMENTS:       N04, S04
GROUP LABEL:              NSTDMAT4          LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                  M811108           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  3
```

Table F-3 (continued)
Conditioning Variables for Mathematics Main Samples

```
001   NSTDMA-A   (1        )  00        NO MORE MATH STUDY (GRADE 4): AGREE
002   NSTDMA-U   (2        )  10        NO MORE MATH STUDY (GRADE 4): UNDECIDED
003   NSTDMA-D   (3,M      )  01        NO MORE MATH STUDY (GRADE 4): DISAGREE, MISSING

CONDITIONING ID:           MATH0030
DESCRIPTION:               AGREE/DISAGREE: I LIKE MATH (GRADES 8, 12)
GRADES/ASSESSMENTS:        N08, S08, N12
GROUP LABEL:               LIKEMATH             LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                   M810701              DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS                NUMBER OF SPECIFICATION RECORDS:  5
001   LKMAT-SA   (1        )  0000      I LIKE MATH (GRADE 8, 12):  STRONGLY AGREE
002   LKMAT-A    (2        )  1000      I LIKE MATH (GRADE 8, 12):  AGREE
003   LKMAT-U    (3        )  0100      I LIKE MATH (GRADE 8, 12):  UNDECIDED
004   LKMAT-D    (4        )  0010      I LIKE MATH (GRADE 8, 12):  DISAGREE
005   LKMAT-SD   (5,M      )  0001      I LIKE MATH (GRADE 8, 12):  STRONGLY DISAGREE, MISSING

CONDITIONING ID:           MATH0031
DESCRIPTION:               AGREE/DISAGREE: I AM GOOD AT MATH (GRADES 8, 12)
GRADES/ASSESSMENTS:        N08, S08, N12
GROUP LABEL:               GOODMATH             LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                   M810703              DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS                NUMBER OF SPECIFICATION RECORDS:  5
001   GDMAT-SA   (1        )  0000      I AM GOOD AT MATH (GRADE 8, 12): STRONGLY AGREE
002   GDMAT-A    (2        )  1000      I AM GOOD AT MATH (GRADE 8, 12): AGREE
003   GDMAT-U    (3        )  0100      I AM GOOD AT MATH (GRADE 8, 12): UNDECIDED
004   GDMAT-D    (4        )  0010      I AM GOOD AT MATH (GRADE 8, 12): DISAGREE
005   GDMAT-SD   (5,M      )  0001      I AM GOOD AT MATH (GRADE 8, 12): STRONGLY DISAGREE, MISSING

CONDITIONING ID:           MATH0032
DESCRIPTION:               AGREE/DISAGREE: IF CHOICE, WOULD NOT STUDY ANY MORE MATH (GRADE 8)
GRADES/ASSESSMENTS:        N08, S08
GROUP LABEL:               NSTDMATH             LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                   M810706              DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS                NUMBER OF SPECIFICATION RECORDS:  5
001   NSMAT-SA   (1,M      )  0000      NO MORE MATH STUDY (GRADE 8): STRONGLY AGREE, MISSING
002   NSMAT-A    (2        )  1000      NO MORE MATH STUDY (GRADE 8): AGREE
003   NSMAT-U    (3        )  0100      NO MORE MATH STUDY (GRADE 8): UNDECIDED
004   NSMAT-D    (4        )  0010      NO MORE MATH STUDY (GRADE 8): DISAGREE
005   NSMAT-SD   (5        )  0001      NO MORE MATH STUDY (GRADE 8): STRONGLY DISAGREE

CONDITIONING ID:           MATH0033
DESCRIPTION:               AGREE/DISAGREE: I UNDERSTAND MOST OF MATH CLASS (GRADE 8, 12)
GRADES/ASSESSMENTS:        N08, S08, N12
GROUP LABEL:               UNDSTMAT             LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                   M810707              DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS                NUMBER OF SPECIFICATION RECORDS:  5
001   USMAT-SA   (1        )  0000      UNDERSTAND MATH (GRADE 8, 12): STRONGLY AGREE
002   USMAT-A    (2        )  1000      UNDERSTAND MATH (GRADE 8, 12): AGREE
003   USMAT-U    (3        )  0100      UNDERSTAND MATH (GRADE 8, 12): UNDECIDED
004   USMAT-D    (4        )  0010      UNDERSTAND MATH (GRADE 8, 12): DISAGREE
005   USMAT-SD   (5,M      )  0001      UNDERSTAND MATH (GRADE 8, 12): STRONGLY DISAGREE, MISSING

CONDITIONING ID:           MATH0034
DESCRIPTION:               AGREE/DISAGREE: MATH IS MORE FOR BOYS THAN FOR GIRLS (GRADE 8, 12)
GRADES/ASSESSMENTS:        N08, S08, N12
GROUP LABEL:               MAT4BOYS             LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                   M810704              DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS                NUMBER OF SPECIFICATION RECORDS:  5
001   M4BOY-SA   (1        )  0000      MATH FOR BOYS (GRADE 8, 12): STRONGLY AGREE
002   M4BOY-A    (2        )  1000      MATH FOR BOYS (GRADE 8, 12): AGREE
003   M4BOY-U    (3        )  0100      MATH FOR BOYS (GRADE 8, 12): UNDECIDED
004   M4BOY-D    (4        )  0010      MATH FOR BOYS (GRADE 8, 12): DISAGREE
005   M4BOY-SD   (5,M      )  0001      MATH FOR BOYS (GRADE 8, 12): STRONGLY DISAGREE, MISSING

CONDITIONING ID:           MATH0035
```

Table F-3 (continued)
Conditioning Variables for Mathematics Main Samples

```
DESCRIPTION:               AGREE/DISAGREE: MATH IS MOSTLY MEMORIZING FACTS (GRADE 8, 12)
GRADES/ASSESSMENTS:        N08, S08, N12
GROUP LABEL:               MATHMEMF            LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                   M810708             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS               NUMBER OF SPECIFICATION RECORDS:  5
001   MATMF-SA   (1          ) 0000            MATH IS MEMORIZING FACTS (GRADE 8, 12): STRONGLY AGREE
002   MATMF-A    (2          ) 1000            MATH IS MEMORIZING FACTS (GRADE 8, 12): AGREE
003   MATMF-U    (3          ) 0100            MATH IS MEMORIZING FACTS (GRADE 8, 12): UNDECIDED
004   MATMF-D    (4          ) 0010            MATH IS MEMORIZING FACTS (GRADE 8, 12): DISAGREE
005   MATMF-SD   (5,M        ) 0001            MATH IS MEMORIZING FACTS (8, 12): STRONGLY DISAGREE, MISSING

CONDITIONING ID:           MATH0036
DESCRIPTION:               AGREE/DISAGREE: ALMOST ALL PEOPLE USE MATH IN THEIR JOBS (GRADE 8, 12)
GRADES/ASSESSMENTS:        N08, S08, N12
GROUP LABEL:               MATHJOBS            LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                   M810702             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS               NUMBER OF SPECIFICATION RECORDS:  5
001   MAJOB-SA   (1          ) 0000            USE MATH EVERYDAY PROBLEMS (GRADE 8, 12): STRONGLY AGREE
002   MAJOB-A    (2          ) 1000            USE MATH IN JOBS (GRADE 8, 12): AGREE
003   MAJOB-U    (3          ) 0100            USE MATH IN JOBS (GRADE 8, 12): UNDECIDED
004   MAJOB-D    (4          ) 0010            USE MATH IN JOBS (GRADE 8, 12): DISAGREE
005   MAJOB-SD   (5,M        ) 0001            USE MATH IN JOBS (GRADE 8, 12): STRONGLY DISAGREE, MISSING

CONDITIONING ID:           MATH0037
DESCRIPTION:               AGREE/DISAGREE: MATH IS USEFUL FOR SOLVING EVERYDAY PROBLEMS (GRADE 8, 12)
GRADES/ASSESSMENTS:        N08, S08, N12
GROUP LABEL:               MATHPROB            LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                   M810705             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS               NUMBER OF SPECIFICATION RECORDS:  5
001   MAPRB-SA   (1          ) 0000            MATH FOR EVERYDAY PROBLEMS (GRADE 8, 12): STRONGLY AGREE
002   MAPRB-A    (2          ) 1000            MATH FOR EVERYDAY PROBLEMS (GRADE 8, 12): AGREE
003   MAPRB-U    (3          ) 0100            MATH FOR EVERYDAY PROBLEMS (GRADE 8, 12): UNDECIDED
004   MAPRB-D    (4          ) 0010            MATH FOR EVERYDAY PROBLEMS (GRADE 8, 12): DISAGREE
005   MAPRB-SD   (5,M        ) 0001            MATH FOR EVERYDAY PROBS (8, 12): STRONGLY DISAGREE, MISSING

CONDITIONING ID:           MATH0038
DESCRIPTION:               HOW LONG HAVE YOU TAKEN GENERAL MATH (9-12)?
GRADES/ASSESSMENTS:        N12
GROUP LABEL:               GEN_MATH            LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                   M811001             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   GENMATH1   (1,M        ) 000             GENERAL MATH:  MORE THAN ONE YEAR, MISSING
002   GENMATH2   (2          ) 100             GENERAL MATH:  ONE SCHOOL YEAR
003   GENMATH3   (3          ) 010             GENERAL MATH:  ONE HALF YEAR OR LESS
004   GENMATH4   (4          ) 001             GENERAL MATH:  NOT STUDIED

CONDITIONING ID:           MATH0039
DESCRIPTION:               HOW LONG HAVE YOU TAKEN BUSINESS/CONSUMER MATH?
GRADES/ASSESSMENTS:        N12
GROUP LABEL:               BUS_MATH            LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                   M811002             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   BUSMATH1   (1,M        ) 000             BUSINESS MATH: MORE THAN ONE YEAR, MISSING
002   BUSMATH2   (2          ) 100             BUSINESS MATH: ONE SCHOOL YEAR
003   BUSMATH3   (3          ) 010             BUSINESS MATH: ONE HALF YEAR OR LESS
004   BUSMATH4   (4          ) 001             BUSINESS MATH: NOT STUDIED

CONDITIONING ID:           MATH0040
DESCRIPTION:               HOW LONG HAVE YOU TAKEN INTRO TO ALGEBRA/PRE-ALGEBRA?
GRADES/ASSESSMENTS:        N12
GROUP LABEL:               PRE-ALGB            LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                   M811003             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   PRE-ALG1   (1,M        ) 000             PRE-ALGEBRA: MORE THAN ONE YEAR, MISSING
002   PRE-ALG2   (2          ) 100             PRE-ALGEBRA: ONE SCHOOL YEAR
```

BEST COPY AVAILABLE

Table F-3 (continued)
Conditioning Variables for Mathematics Main Samples

```
003   PRE-ALG3   (3          )   010          PRE-ALGEBRA: ONE HALF YEAR OR LESS
004   PRE-ALG4   (4          )   001          PRE-ALGEBRA: NOT STUDIED

CONDITIONING ID:           MATH0041
DESCRIPTION:               HOW LONG HAVE YOU TAKEN ALGEBRA 1?
GRADES/ASSESSMENTS:        N12
GROUP LABEL:               ALGBRA_1            LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                   M811004             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS               NUMBER OF SPECIFICATION RECORDS:   4
001   ALGBR1-1   (1          )   000          ALGEBRA 1: MORE THAN ONE YEAR
002   ALGBR1-2   (2          )   100          ALGEBRA 1: ONE SCHOOL YEAR
003   ALGBR1-3   (3          )   010          ALGEBRA 1: ONE HALF YEAR OR LESS
004   ALGBR1-4   (4,M        )   001          ALGEBRA 1: NOT STUDIED, MISSING

CONDITIONING ID:           MATH0042
DESCRIPTION:               HOW LONG HAVE YOU TAKEN GEOMETRY?
GRADES/ASSESSMENTS:        N12
GROUP LABEL:               GEOMETRY            LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                   M811005             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS               NUMBER OF SPECIFICATION RECORDS:   4
001   GEOMTRY1   (1          )   000          GEOMETRY: MORE THAN ONE YEAR
002   GEOMTRY2   (2          )   100          GEOMETRY: ONE SCHOOL YEAR
003   GEOMTRY3   (3          )   010          GEOMETRY: ONE HALF YEAR OR LESS
004   GEOMTRY4   (4,M        )   001          GEOMETRY: NOT STUDIED, MISSING

CONDITIONING ID:           MATH0043
DESCRIPTION:               HOW LONG HAVE YOU TAKEN ALGEBRA 2?
GRADES/ASSESSMENTS:        N12
GROUP LABEL:               ALGBRA_2            LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                   M811006             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS               NUMBER OF SPECIFICATION RECORDS:   4
001   ALGBR2-1   (1          )   000          ALGEBRA 2: MORE THAN ONE YEAR
002   ALGBR2-2   (2          )   100          ALGEBRA 2: ONE SCHOOL YEAR
003   ALGBR2-3   (3          )   010          ALGEBRA 2: ONE HALF YEAR OR LESS
004   ALGBR2-4   (4,M        )   001          ALGEBRA 2: NOT STUDIED, MISSING

CONDITIONING ID:           MATH0044
DESCRIPTION:               HOW LONG HAVE YOU TAKEN TRIGONOMETRY?
GRADES/ASSESSMENTS:        N12
GROUP LABEL:               TRIGMTRY            LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                   M811007             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS               NUMBER OF SPECIFICATION RECORDS:   4
001   TRIGTRY1   (1          )   000          TRIGONOMETRY: MORE THAN ONE YEAR
002   TRIGTRY2   (2          )   100          TRIGONOMETRY: ONE SCHOOL YEAR
003   TRIGTRY3   (3          )   010          TRIGONOMETRY: ONE HALF YEAR OR LESS
004   TRIGTRY4   (4,M        )   001          TRIGONOMETRY: NOT STUDIED, MISSING

CONDITIONING ID:           MATH0045
DESCRIPTION:               HOW LONG HAVE YOU TAKEN PRE-CALCULUS?
GRADES/ASSESSMENTS:        N12
GROUP LABEL:               PRE-CALC            LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                   M811008             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS               NUMBER OF SPECIFICATION RECORDS:   4
001   PRE-CAL1   (1          )   000          PRE-CALCULUS: MORE THAN ONE YEAR
002   PRE-CAL2   (2          )   100          PRE-CALCULUS: ONE SCHOOL YEAR
003   PRE-CAL3   (3          )   010          PRE-CALCULUS: ONE HALF YEAR OR LESS
004   PRE-CAL4   (4,M        )   001          PRE-CALCULUS: NOT STUDIED, MISSING

CONDITIONING ID:           MATH0046
DESCRIPTION:               HOW LONG HAVE YOU TAKEN PROBABILITY/STATISTICS?
GRADES/ASSESSMENTS:        N12
GROUP LABEL:               PROBSTAT            LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                   M811009             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS               NUMBER OF SPECIFICATION RECORDS:   4
001   PROBSTA1   (1          )   000          PROBABILITY/STATISTICS: MORE THAN ONE YEAR
```

Table F-3 (continued)
Conditioning Variables for Mathematics Main Samples

```
002   PROBSTA2   (2          )  100        PROBABILITY/STATISTICS: ONE SCHOOL YEAR
003   PROBSTA3   (3          )  010        PROBABILITY/STATISTICS: ONE HALF YEAR OR LESS
004   PROBSTA4   (4,M        )  001        PROBABILITY/STATISTICS: NOT STUDIED, MISSING

CONDITIONING ID:             MATH0047
DESCRIPTION:                 HOW LONG HAVE YOU TAKEN UNIFIED/INTEGRATED/SEQUENTIAL MATH?
GRADES/ASSESSMENTS:          N12
GROUP LABEL:                 UNIF_MAT          LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                     M811012           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS             NUMBER OF SPECIFICATION RECORDS:  4
001   UNIFMAT1   (1          )  000        UNIFIED MATH: MORE THAN ONE YEAR
002   UNIFMAT2   (2          )  100        UNIFIED MATH: ONE SCHOOL YEAR
003   UNIFMAT3   (3          )  010        UNIFIED MATH: ONE HALF YEAR OR LESS
004   UNIFMAT4   (4,M        )  001        UNIFIED MATH: NOT STUDIED, MISSING

CONDITIONING ID:             MATH0048
DESCRIPTION:                 HOW LONG HAVE YOU TAKEN CALCULUS?
GRADES/ASSESSMENTS:          N12
GROUP LABEL:                 CALCULUS          LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                     M811011           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS             NUMBER OF SPECIFICATION RECORDS:  4
001   CALCLUS1   (1          )  000        CALCULUS: MORE THAN ONE YEAR
002   CALCLUS2   (2          )  100        CALCULUS: ONE SCHOOL YEAR
003   CALCLUS3   (3          )  010        CALCULUS: ONE HALF YEAR OR LESS
004   CALCLUS4   (4,M        )  001        CALCULUS: NOT STUDIED, MISSING

CONDITIONING ID:             MATH0049
DESCRIPTION:                 HOW LONG HAVE YOU TAKEN COMPUTER SCIENCE?
GRADES/ASSESSMENTS:          N12
GROUP LABEL:                 COMP_SCI          LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                     M811013           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS             NUMBER OF SPECIFICATION RECORDS:  4
001   COMPSCI1   (1          )  000        COMPUTER SCIENCE: MORE THAN ONE YEAR
002   COMPSCI2   (2          )  100        COMPUTER SCIENCE: ONE SCHOOL YEAR
003   COMPSCI3   (3          )  010        COMPUTER SCIENCE: ONE HALF YEAR OR LESS
004   COMPSCI4   (4,M        )  001        COMPUTER SCIENCE: NOT STUDIED, MISSING

CONDITIONING ID:             MATH0050
DESCRIPTION:                 ABOUT HOW MANY QUESTIONS DID YOU GET RIGHT?
GRADES/ASSESSMENTS:          N04, S04, N08, S08, N12
GROUP LABEL:                 #QUESTN+          LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                     MM00101           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS             NUMBER OF SPECIFICATION RECORDS:  4
001   #QUEST+1   (1          )  000        NUMBER QUESTIONS RIGHT:  ALMOST ALL
002   #QUEST+2   (2          )  100        NUMBER QUESTIONS RIGHT:  MORE THAN HALF
003   #QUEST+3   (3          )  010        NUMBER QUESTIONS RIGHT:  ABOUT HALF
004   #QUEST+4   (4,M        )  001        NUMBER QUESTIONS RIGHT:  LESS THAN HALF, MISSING

CONDITIONING ID:             MATH0051
DESCRIPTION:                 HOW HARD WAS THIS TEST COMPARED TO OTHERS?
GRADES/ASSESSMENTS:          N04, S04, N08, S08, N12
GROUP LABEL:                 TEST_DIF          LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                     MM00201           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS             NUMBER OF SPECIFICATION RECORDS:  4
001   TESTDIF1   (1,M        )  000        TEST DIFFICULTY:  MUCH HARDER THAN OTHERS
002   TESTDIF2   (2          )  100        TEST DIFFICULTY:  HARDER THAN OTHERS
003   TESTDIF3   (3          )  010        TEST DIFFICULTY:  ABOUT AS HARD AS OTHERS
004   TESTDIF4   (4          )  001        TEST DIFFICULTY:  EASIER THAN OTHERS

CONDITIONING ID:             MATH0052
DESCRIPTION:                 HOW HARD DID YOU TRY ON THIS TEST COMPRAED TO OTHERS?
GRADES/ASSESSMENTS:          N04, S04, N08, S08, N12
GROUP LABEL:                 TEST_EFF          LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                     MM00301           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS             NUMBER OF SPECIFICATION RECORDS:  4
```

Table F-3 (continued)
Conditioning Variables for Mathematics Main Samples

```
001   TESTEFF1   (1,M         )  000         TEST EFFORT:  MUCH HARDER THAN OTHERS
002   TESTEFF2   (2           )  100         TEST EFFORT:  HARDER THAN OTHERS
003   TESTEFF3   (3           )  010         TEST EFFORT:  ABOUT AS HARD AS OTHERS
004   TESTEFF4   (4           )  001         TEST EFFORT:  NOT AS HARD AS OTHERS

CONDITIONING ID:            MATH0053
DESCRIPTION:                HOW IMPORTANT WAS IT TO YOU TO DO WELL?
GRADES/ASSESSMENTS:         N04, S04, N08, S08, N12
GROUP LABEL:                TEST_IMP          LENGTH OF CONTRAST FIELD        :   4
NAEP ID:                    MM00401           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS             NUMBER OF SPECIFICATION RECORDS:   5
001   TESTIMP1   (1           )  0000        TEST IMPORTANCE:  VERY IMPORTANT
002   TESTIMP2   (2           )  1000        TEST IMPORTANCE:  IMPORTANT
003   TESTIMP3   (3           )  0100        TEST IMPORTANCE:  SOMEWHAT IMPORTANT
004   TESTIMP4   (4           )  0010        TEST IMPORTANCE:  NOT VERY IMPORTANT
005   TESTIMP?   (M           )  0001        TEST IMPORTANCE:  MISSING

CONDITIONING ID:            MATH0054
DESCRIPTION:                HOW OFTEN WERE YOU ASKED TO PROVIDE DETAILED SOLUTIONS ON TESTS?
GRADES/ASSESSMENTS:         N04, S04, N08, S08, N12
GROUP LABEL:                DSOLUTNS          LENGTH OF CONTRAST FIELD        :   3
NAEP ID:                    MM00501           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS             NUMBER OF SPECIFICATION RECORDS:   4
001   DSOLUTN1   (1           )  000         DETAILED SOLUTIONS:  AT LEAST ONCE A WEEK
002   DSOLUTN2   (2           )  100         DETAILED SOLUTIONS:  ONCE OR TWICE A MONTH
003   DSOLUTN3   (3           )  010         DETAILED SOLUTIONS:  ONCE OR TWICE A YEAR
004   DSOLUTN4   (4,M         )  001         DETAILED SOLUTIONS:  NEVER, MISSING

CONDITIONING ID:            SCHL0001
DESCRIPTION:                ARE TRIGONOMETRY COURSES TAUGHT AT YOUR SCHOOL?
GRADES/ASSESSMENTS:         N12
GROUP LABEL:                TRIG@SCH          LENGTH OF CONTRAST FIELD        :   2
NAEP ID:                    C035307           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS             NUMBER OF SPECIFICATION RECORDS:   3
001   TRISCH-Y   (1           )  00          TRIGONOMETRY TAUGHT?:  YES
002   TRISCH-N   (2           )  10          TRIGONOMETRY TAUGHT?:  NO
003   TRISCH-?   (M           )  01          TRIGONOMETRY TAUGHT?:  MISSING

CONDITIONING ID:            SCHL0002
DESCRIPTION:                ARE PRE-CALCULUS COURSES TAUGHT AT YOUR SCHOOL?
GRADES/ASSESSMENTS:         N12
GROUP LABEL:                PCAL@SCH          LENGTH OF CONTRAST FIELD        :   2
NAEP ID:                    C035308           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS             NUMBER OF SPECIFICATION RECORDS:   3
001   PCASCH-Y   (1           )  00          PRE-CALCULUS TAUGHT?:  YES
002   PCASCH-N   (2           )  10          PRE-CALCULUS TAUGHT?:  NO
003   PCASCH-?   (M           )  01          PRE-CALCULUS TAUGHT?:  MISSING

CONDITIONING ID:            SCHL0003
DESCRIPTION:                ARE PROBABILITY/STATISTICS COURSES TAUGHT AT YOUR SCHOOL?
GRADES/ASSESSMENTS:         N12
GROUP LABEL:                P/S@SCH           LENGTH OF CONTRAST FIELD        :   2
NAEP ID:                    C035310           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS             NUMBER OF SPECIFICATION RECORDS:   3
001   P/SSCH-Y   (1           )  00          PROB/STAT TAUGHT?:  YES
002   P/SSCH-N   (2           )  10          PROB/STAT TAUGHT?:  NO
003   P/SSCH-?   (M           )  01          PROB/STAT TAUGHT?:  MISSING

CONDITIONING ID:            SCHL0004
DESCRIPTION:                ARE CALCULUS COURSES TAUGHT AT YOUR SCHOOL?
GRADES/ASSESSMENTS:         N12
GROUP LABEL:                CALC@SCH          LENGTH OF CONTRAST FIELD        :   2
NAEP ID:                    C035312           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS             NUMBER OF SPECIFICATION RECORDS:   3
001   CALSCH-Y   (1           )  00          CALCULUS TAUGHT?:  YES
```

Table F-3 (continued)
Conditioning Variables for Mathematics Main Samples

```
002  CALSCH-N  (2           )  10           CALCULUS TAUGHT?:  NO
003  CALSCH-?  (M           )  01           CALCULUS TAUGHT?:  MISSING

CONDITIONING ID:             SCHL0005
DESCRIPTION:                 ARE COMPUTER SCIENCE COURSES TAUGHT AT YOUR SCHOOL?
GRADES/ASSESSMENTS:          N12
GROUP LABEL:                 COMP@SCH           LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                     C035313            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS              NUMBER OF SPECIFICATION RECORDS:  3
001  CMPSCH-Y  (1           )  00           COMPUTER SCIENCE TAUGHT?: YES
002  CMPSCH-N  (2           )  10           COMPUTER SCIENCE TAUGHT?: NO
003  CMPSCH-?  (M           )  01           COMPUTER SCIENCE TAUGHT?: MISSING

CONDITIONING ID:             SCHL0006
DESCRIPTION:                 DOES SCHOOL OFFER ALGEBRA TO 8TH GRADE FOR HS CREDIT?
GRADES/ASSESSMENTS:          N08, S08
GROUP LABEL:                 ALG4HSCR           LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                     C034600            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS              NUMBER OF SPECIFICATION RECORDS:  3
001  ALG4HS-Y  (1           )  00           ALGEBRA FOR HS CREDIT:  YES
002  ALG4HS-N  (2           )  10           ALGEBRA FOR HS CREDIT:  NO
003  ALG4HS-?  (M           )  01           ALGEBRA FOR HS CREDIT:  MISSING

CONDITIONING ID:             SCHL0007
DESCRIPTION:                 WHO TEACHES ENGLISH/LANGUAGE ARTS TO 8TH GRADE?
GRADES/ASSESSMENTS:          N08, S08
GROUP LABEL:                 TSUB_ENG           LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                     C034701            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS              NUMBER OF SPECIFICATION RECORDS:  3
001  ENG-T>1S  (1,M         )  00           8TH-GRADE ENGLISH:  TEACHERS WITH MORE THAN ONE SUBJECT
002  ENG-T=1S  (2           )  10           8TH-GRADE ENGLISH:  TEACHERS WITH ONE SUBJECT
003  ENG-SNT   (3           )  01           8TH-GRADE ENGLISH:  SUBJECT NOT TAUGHT

CONDITIONING ID:             SCHL0008
DESCRIPTION:                 WHO TEACHES MATHEMATICS TO 8TH GRADE?
GRADES/ASSESSMENTS:          N08, S08
GROUP LABEL:                 TSUB_MAT           LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                     C034702            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS              NUMBER OF SPECIFICATION RECORDS:  3
001  MAT-T>1S  (1,M         )  00           8TH-GRADE MATH:  TEACHERS WITH MORE THAN ONE SUBJECT
002  MAT-T=1S  (2           )  10           8TH-GRADE MATH:  TEACHERS WITH ONE SUBJECT
003  MAT-SNT   (3           )  01           8TH-GRADE MATH:  SUBJECT NOT TAUGHT

CONDITIONING ID:             SCHL0009
DESCRIPTION:                 HAS READING BEEN IDENTIFIED AS A PRIORITY? (GRADE 4)
GRADES/ASSESSMENTS:          N04, S04, N08, S08
GROUP LABEL:                 PRIOR-RD           LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                     C031601            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS              NUMBER OF SPECIFICATION RECORDS:  3
001  RPRIOR-Y  (1           )  00           READING PRIORITY:  YES
002  RPRIOR-N  (2           )  10           READING PRIORITY:  NO
003  RPRIOR-?  (M           )  01           READING PRIORITY:  MISSING

CONDITIONING ID:             SCHL0010
DESCRIPTION:                 HAS WRITING BEEN IDENTIFIED AS A PRIORITY? (GRADE 4)
GRADES/ASSESSMENTS:          N04, S04, N08, S08
GROUP LABEL:                 PRIOR-WR           LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                     C031602            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS              NUMBER OF SPECIFICATION RECORDS:  3
001  WPRIOR-Y  (1           )  00           WRITING PRIORITY:  YES
002  WPRIOR-N  (2           )  10           WRITING PRIORITY:  NO
003  WPRIOR-?  (M           )  01           WRITING PRIORITY:  MISSING

CONDITIONING ID:             SCHL0011
DESCRIPTION:                 HAS MATHEMATICS BEEN IDENTIFIED AS A PRIORITY? (GRADE 4)
```

Table F-3 (continued)
Conditioning Variables for Mathematics Main Samples

```
GRADES/ASSESSMENTS:        N04, S04, N08, S08
GROUP LABEL:               PRIOR-MA          LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                   C031603           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:  3
001   MPRIOR-Y   (1          )  00           MATH PRIORITY:  YES
002   MPRIOR-N   (2          )  10           MATH PRIORITY:  NO
003   MPRIOR-?   (M          )  01           MATH PRIORITY:  MISSING

CONDITIONING ID:           SCHL0012
DESCRIPTION:               WHAT PERCENT OF STUDENTS RECEIVE SUBSIDIZED LUNCH?
GRADES/ASSESSMENTS:        N04, S04, N08, S08, N12
GROUP LABEL:               %SUBLUN           LENGTH OF CONTRAST FIELD       :  8
NAEP ID:                   C032001           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:  9
001   %SUBLUN1   (1          )  00000000     PERCENT SUBSIDIZED LUNCH:  NONE
002   %SUBLUN2   (2          )  10000000     PERCENT SUBSIDIZED LUNCH:  1-5%
003   %SUBLUN3   (3          )  01000000     PERCENT SUBSIDIZED LUNCH:  6-10%
004   %SUBLUN4   (4          )  00100000     PERCENT SUBSIDIZED LUNCH:  11-25%
005   %SUBLUN5   (5          )  00010000     PERCENT SUBSIDIZED LUNCH:  26-50%
006   %SUBLUN6   (6          )  00001000     PERCENT SUBSIDIZED LUNCH:  51-75%
007   %SUBLUN7   (7          )  00000100     PERCENT SUBSIDIZED LUNCH:  76-90%
008   %SUBLUN8   (8          )  00000010     PERCENT SUBSIDIZED LUNCH:  90-100%
009   %SUBLUN?   (M          )  00000001     PERCENT SUBSIDIZED LUNCH:  MISSING

CONDITIONING ID:           SCHL0013
DESCRIPTION:               WHAT PERCENT OF STUDENTS RECEIVE REMEDIAL READING?
GRADES/ASSESSMENTS:        N04, S04, N08, S08, N12
GROUP LABEL:               %REMDL-R          LENGTH OF CONTRAST FIELD       :  8
NAEP ID:                   C032002           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:  9
001   %REMRED1   (1          )  00000000     PERCENT REMEDIAL READING:  NONE
002   %REMRED2   (2          )  10000000     PERCENT REMEDIAL READING:  1-5%
003   %REMRED3   (3          )  01000000     PERCENT REMEDIAL READING:  6-10%
004   %REMRED4   (4          )  00100000     PERCENT REMEDIAL READING:  11-25%
005   %REMRED5   (5          )  00010000     PERCENT REMEDIAL READING:  26-50%
006   %REMRED6   (6          )  00001000     PERCENT REMEDIAL READING:  51-75%
007   %REMRED7   (7          )  00000100     PERCENT REMEDIAL READING:  76-90%
008   %REMRED8   (8          )  00000010     PERCENT REMEDIAL READING:  90-100%
009   %REMRED?   (M          )  00000001     PERCENT REMEDIAL READING:  MISSING

CONDITIONING ID:           SCHL0014
DESCRIPTION:               WHAT PERCENT OF STUDENTS RECEIVE REMEDIAL MATHEMATICS?
GRADES/ASSESSMENTS:        N04, S04, N08, S08, N12
GROUP LABEL:               %REMDL-M          LENGTH OF CONTRAST FIELD       :  8
NAEP ID:                   C032003           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:  9
001   %REMMAT1   (1          )  00000000     PERCENT REMEDIAL MATH:  NONE
002   %REMMAT2   (2          )  10000000     PERCENT REMEDIAL MATH:  1-5%
003   %REMMAT3   (3          )  01000000     PERCENT REMEDIAL MATH:  6-10%
004   %REMMAT4   (4          )  00100000     PERCENT REMEDIAL MATH:  11-25%
005   %REMMAT5   (5          )  00010000     PERCENT REMEDIAL MATH:  26-50%
006   %REMMAT6   (6          )  00001000     PERCENT REMEDIAL MATH:  51-75%
007   %REMMAT7   (7          )  00000100     PERCENT REMEDIAL MATH:  76-90%
008   %REMMAT8   (8          )  00000010     PERCENT REMEDIAL MATH:  90-100%
009   %REMMAT?   (M          )  00000001     PERCENT REMEDIAL MATH:  MISSING

CONDITIONING ID:           SCHL0015
DESCRIPTION:               WHAT PERCENTAGE OF STUDENTS ARE ENROLLED AT BEGINNING AND END OF SCHOOL YEAR?
GRADES/ASSESSMENTS:        N04, S04, N08, S08, N12
GROUP LABEL:               %ENR/YR           LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                   C033700           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:  5
001   %ENR/YR1   (1          )  0000         YEAR LONG ENROLLMENT:  98-100 PERCENT
002   %ENR/YR2   (2          )  1000         YEAR LONG ENROLLMENT:  95-97 PERCENT
003   %ENR/YR3   (3          )  0100         YEAR LONG ENROLLMENT:  90-94 PERCENT
```

Table F-3 (continued)
Conditioning Variables for Mathematics Main Samples

```
004   %ENR/YR4   (4          )  0010        YEAR LONG ENROLLMENT:  LESS THAN 90 PERCENT
005   %ENR/YR?   (M          )  0001        YEAR LONG ENROLLMENT:  MISSING

CONDITIONING ID:            SCHL0016
DESCRIPTION:                WHAT PERCENTAGE OF 4TH GRADERS RETAINED IN 91-92?
GRADES/ASSESSMENTS:         N04, S04, N08, S08
GROUP LABEL:                %4RETAIN          LENGTH OF CONTRAST FIELD       :  5
NAEP ID:                    C033800           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS             NUMBER OF SPECIFICATION RECORDS:  6
001   %RETAIN1   (1          )  00000       % 4TH GRADE RETAINED:  0 PERCENT
002   %RETAIN2   (2          )  10000       % 4TH GRADE RETAINED:  1-2 PERCENT
003   %RETAIN3   (3          )  01000       % 4TH GRADE RETAINED:  3-5 PERCENT
004   %RETAIN4   (4          )  00100       % 4TH GRADE RETAINED:  6-10 PERCENT
005   %RETAIN5   (5          )  00010       % 4TH GRADE RETAINED:  MORE THAN 10 PERCENT
006   %RETAIN?   (M          )  00001       % 4TH GRADE RETAINED:  MISSING

CONDITIONING ID:            SCHL0017
DESCRIPTION:                WHAT PERCENTAGE OF TEACHERS NOT AT SCHOOL AT END OF YEAR?
GRADES/ASSESSMENTS:         N04, S04, N08, S08, N12
GROUP LABEL:                %T_LEAVE          LENGTH OF CONTRAST FIELD       :  5
NAEP ID:                    C033903           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS             NUMBER OF SPECIFICATION RECORDS:  6
001   %TLEAVE1   (1          )  00000       % TEACHERS LEAVE/YEAR: 0 PERCENT
002   %TLEAVE2   (2          )  10000       % TEACHERS LEAVE/YEAR: 1-2 PERCENT
003   %TLEAVE3   (3          )  01000       % TEACHERS LEAVE/YEAR: 3-5 PERCENT
004   %TLEAVE4   (4          )  00100       % TEACHERS LEAVE/YEAR: 6-10 PERCENT
005   %TLEAVE5   (5          )  00010       % TEACHERS LEAVE/YEAR: MORE THAN 10 PERCENT
006   %TLEAVE?   (M          )  00001       % TEACHERS LEAVE/YEAR: MISSING

CONDITIONING ID:            SCHL0018
DESCRIPTION:                PERCENTAGE OF GRADUATING CLASS IN TWO-YEAR COLLEGES
GRADES/ASSESSMENTS:         N12
GROUP LABEL:                %2YR_COL          LENGTH OF CONTRAST FIELD       :  6
NAEP ID:                    C036001           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS             NUMBER OF SPECIFICATION RECORDS:  7
001   %2YRCOL1   (1          )  000000      2-YEAR COLLEGE:  0-10 PERCENT
002   %2YRCOL2   (2          )  100000      2-YEAR COLLEGE:  11-25 PERCENT
003   %2YRCOL3   (3          )  010000      2-YEAR COLLEGE:  26-50 PERCENT
004   %2YRCOL4   (4          )  001000      2-YEAR COLLEGE:  51-75 PERCENT
005   %2YRCOL5   (5          )  000100      2-YEAR COLLEGE:  76-90 PERCENT
006   %2YRCOL6   (6          )  000010      2-YEAR COLLEGE:  91-100 PERCENT
007   %2YRCOL?   (M          )  000001      2-YEAR COLLEGE:  MISSING

CONDITIONING ID:            SCHL0019
DESCRIPTION:                PERCENTAGE OF GRADUATING CLASS IN FOUR-YEAR COLLEGES
GRADES/ASSESSMENTS:         N12
GROUP LABEL:                %4YR_COL          LENGTH OF CONTRAST FIELD       :  6
NAEP ID:                    C036002           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS             NUMBER OF SPECIFICATION RECORDS:  7
001   %4YRCOL1   (1          )  000000      4-YEAR COLLEGE:  0-10 PERCENT
002   %4YRCOL2   (2          )  100000      4-YEAR COLLEGE:  11-25 PERCENT
003   %4YRCOL3   (3          )  010000      4-YEAR COLLEGE:  26-50 PERCENT
004   %4YRCOL4   (4          )  001000      4-YEAR COLLEGE:  51-75 PERCENT
005   %4YRCOL5   (5          )  000100      4-YEAR COLLEGE:  76-90 PERCENT
006   %4YRCOL6   (6          )  000010      4-YEAR COLLEGE:  91-100 PERCENT
007   %4YRCOL?   (M          )  000001      4-YEAR COLLEGE:  MISSING

CONDITIONING ID:            TCHR0001
DESCRIPTION:                HOW WELL DOES SCHOOL PROVIDE RESOURCES
GRADES/ASSESSMENTS:         N04, S04, N08, S08
GROUP LABEL:                RESOURCE          LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                    T041201           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS             NUMBER OF SPECIFICATION RECORDS:  5
001   RESOURC1   (1          )  0000        RESOURCES:  GET ALL
002   RESOURC2   (2          )  1000        RESOURCES:  GET MOST
```

Table F-3 (continued)
Conditioning Variables for Mathematics Main Samples

```
003    RESOURC3    (3           )  0100        RESOURCES:  GET SOME
004    RESOURC4    (4           )  0010        RESOURCES:  DON'T GET
005    RESOURC?    (M,DNA       )  0001        RESOURCES:  MISSING, DOES NOT APPLY

CONDITIONING ID:               TCHR0002
DESCRIPTION:                   TEACHER MATCH STATUS WITH STUDENT
GRADES/ASSESSMENTS:            N04, S04, N08, S08
GROUP LABEL:                   T_MATCH              LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                       TCHMTCH              DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:              CLASS                NUMBER OF SPECIFICATION RECORDS:  3
001    TMCH-NO     (1,M         )  00          TEACHER MATCH:  NO MATCH
002    TMCH-PAR    (2           )  10          TEACHER MATCH:  PARTIAL MATCH
003    TMCH-COM    (3           )  01          TEACHER MATCH:  COMPLETE MATCH

CONDITIONING ID:               TMAT0001
DESCRIPTION:                   WHAT IS THE MATH ABILITY OF STUDENTS IN THIS CLASS?
GRADES/ASSESSMENTS:            N04, S04, N08, S08
GROUP LABEL:                   ABIL_MAT             LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                       T044100              DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:              CLASS                NUMBER OF SPECIFICATION RECORDS:  5
001    AB_MATH1    (1           )  0000        MATH ABILITY:  PRIMARILY HIGH ABILITY
002    AB_MATH2    (2           )  1000        MATH ABILITY:  PRIMARILY LOW ABILITY
003    AB_MATH3    (3           )  0100        MATH ABILITY:  PRIMARILY AVERAGE ABILITY
004    AB_MATH4    (4           )  0010        MATH ABILITY:  WIDELY MIXED ABILITY
005    AB_MATH?    (M,DNA       )  0001        MATH ABILITY:  MISSING, DOES NOT APPLY

CONDITIONING ID:               TMAT0002
DESCRIPTION:                   AMOUNT OF MATH HOMEWORK ASSIGNED PER DAY (TEACHER)
GRADES/ASSESSMENTS:            N04, S04, N08, S08
GROUP LABEL:                   T_MATHHW             LENGTH OF CONTRAST FIELD       :  6
NAEP ID:                       T044400              DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:              CLASS                NUMBER OF SPECIFICATION RECORDS:  7
001    T_MATHW1    (1           )  000000      AMOUNT MATH HOMEWORK (TEACHER):  NONE
002    T_MATHW2    (2           )  100000      AMOUNT MATH HOMEWORK (TEACHER):  15 MINUTES
003    T_MATHW3    (3           )  010000      AMOUNT MATH HOMEWORK (TEACHER):  30 MINUTES
004    T_MATHW4    (4           )  001000      AMOUNT MATH HOMEWORK (TEACHER):  45 MINUTES
005    T_MATHW5    (5           )  000100      AMOUNT MATH HOMEWORK (TEACHER):  1 HOUR
006    T_MATHW6    (6           )  000010      AMOUNT MATH HOMEWORK (TEACHER):  MORE THAN 1 HOUR
007    T_MATHW?    (M,DNA       )  000001      AMOUNT MATH HOMEWORK (TEACHER):  MISSING, DOES NOT APPLY

CONDITIONING ID:               TMAT0003
DESCRIPTION:                   HOW OFTEN DO STUDENTS DO MATH FROM TEXTBOOKS?
GRADES/ASSESSMENTS:            N04, S04, N08, S08
GROUP LABEL:                   T_TXTBKS             LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                       T044501              DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:              CLASS                NUMBER OF SPECIFICATION RECORDS:  5
001    T_TXTBK1    (1           )  0000        MATH FROM TEXTBOOKS (TEACHER):  ALMOST EVERY DAY
002    T_TXTBK2    (2           )  1000        MATH FROM TEXTBOOKS (TEACHER):  ONCE OR TWICE A WEEK
003    T_TXTBK3    (3           )  0100        MATH FROM TEXTBOOKS (TEACHER):  ONCE OR TWICE A MONTH
004    T_TXTBK4    (4           )  0010        MATH FROM TEXTBOOKS (TEACHER):  NEVER OR HARDLY EVER
005    T_TXTBK?    (M,DNA       )  0001        MATH FROM TEXTBOOKS (TEACHER):  MISSING, DOES NOT APPLY

CONDITIONING ID:               TMAT0004
DESCRIPTION:                   HOW OFTEN DO STUDENTS DO MATH FROM WORKSHEETS?
GRADES/ASSESSMENTS:            N04, S04, N08, S08
GROUP LABEL:                   T_WRKSHS             LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                       T044502              DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:              CLASS                NUMBER OF SPECIFICATION RECORDS:  5
001    T_WRKSH1    (1           )  0000        MATH FROM WORKSHEETS (TEACHER): ALMOST EVERY DAY
002    T_WRKSH2    (2           )  1000        MATH FROM WORKSHEETS (TEACHER): ONCE OR TWICE A WEEK
003    T_WRKSH3    (3           )  0100        MATH FROM WORKSHEETS (TEACHER): ONCE OR TWICE A MONTH
004    T_WRKSH4    (4           )  0010        MATH FROM WORKSHEETS (TEACHER): NEVER OR HARDLY EVER
005    T_WRKSH?    (M,DNA       )  0001        MATH FROM WORKSHEETS (TEACHER): MISSING, DOES NOT APPLY

CONDITIONING ID:               TMAT0005
```

Table F-3 (continued)
Conditioning Variables for Mathematics Main Samples

```
DESCRIPTION:                 HOW OFTEN DO STUDENTS DO MATH IN SMALL GROUPS?
GRADES/ASSESSMENTS:          N04, S04, N08, S08
GROUP LABEL:                 T_SMGRPS          LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                     T044503           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS             NUMBER OF SPECIFICATION RECORDS:  5
001   T_SMGRP1   (1             )  0000        MATH IN SMALL GROUPS (TEACHER): ALMOST EVERY DAY
002   T_SMGRP2   (2             )  1000        MATH IN SMALL GROUPS (TEACHER): ONCE OR TWICE A WEEK
003   T_SMGRP3   (3             )  0100        MATH IN SMALL GROUPS (TEACHER): ONCE OR TWICE A MONTH
004   T_SMGRP4   (4             )  0010        MATH IN SMALL GROUPS (TEACHER): NEVER OR HARDLY EVER
005   T_SMGRP?   (M,DNA         )  0001        MATH IN SMALL GROUPS (TEACHER): MISSING, DOES NOT APPLY

CONDITIONING ID:             TMAT0006
DESCRIPTION:                 HOW OFTEN DO STUDENTS WORK WITH OBJECTS?
GRADES/ASSESSMENTS:          N04, S04
GROUP LABEL:                 T_OBJCTS          LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                     T044504           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS             NUMBER OF SPECIFICATION RECORDS:  5
001   T_OBJCT1   (1             )  0000        WORK WITH OBJECTS (TEACHER): ALMOST EVERY DAY
002   T_OBJCT2   (2             )  1000        WORK WITH OBJECTS (TEACHER): ONCE OR TWICE A WEEK
003   T_OBJCT3   (3             )  0100        WORK WITH OBJECTS (TEACHER): ONCE OR TWICE A MONTH
004   T_OBJCT4   (4             )  0010        WORK WITH OBJECTS (TEACHER): NEVER OR HARDLY EVER
005   T_OBJCT?   (M,DNA         )  0001        WORK WITH OBJECTS (TEACHER): MISSING, DOES NOT APPLY

CONDITIONING ID:             TMAT0007
DESCRIPTION:                 HOW OFTEN STUDENTS WORK WITH MEASUREMENT INSTRUMENTS/GEOMETRIC SOLIDS (TEACHER)?
GRADES/ASSESSMENTS:          N08, S08
GROUP LABEL:                 T_MI&GS           LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                     T044512           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS             NUMBER OF SPECIFICATION RECORDS:  5
001   T_MI&GS1   (1             )  0000        MEASUREMENT INSTRUMENTS/GEOM SOLIDS (TCHR): ALMOST EVERY DAY
002   T_MI&GS2   (2             )  1000        MEASUREMENT INSTRUMENTS/GEOM SOLIDS (TCHR): ONCE OR TWICE/WK
003   T_MI&GS3   (3             )  0100        MEASURE INSTRUMENTS/GEOM SOLIDS (TCHR): ONCE OR TWICE/MNTH
004   T_MI&GS4   (4             )  0010        MEASURE INSTRUMENTS/GEOM SOLIDS (TCHR): NEVER OR HARDLY EVER
005   T_MI&GS?   (M,DNA         )  0001        MEASURE INSTRUMENTS/GEOM SOLIDS (TCHR): MISSING, NOT APPLY

CONDITIONING ID:             TMAT0008
DESCRIPTION:                 HOW OFTEN DO STUDENTS USE A CALCULATOR (TEACHER)?
GRADES/ASSESSMENTS:          N04, S04, N08, S08
GROUP LABEL:                 T_CALCTR          LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                     T044505           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS             NUMBER OF SPECIFICATION RECORDS:  5
001   T_CALCR1   (1             )  0000        USE CALCULATOR (TEACHER): ALMOST EVERY DAY
002   T_CALCR2   (2             )  1000        USE CALCULATOR (TEACHER): ONCE OR TWICE A WEEK
003   T_CALCR3   (3             )  0100        USE CALCULATOR (TEACHER): ONCE OR TWICE A MONTH
004   T_CALCR4   (4             )  0010        USE CALCULATOR (TEACHER): NEVER OR HARDLY EVER
005   T_CALCR?   (M,DNA         )  0001        USE CALCULATOR (TEACHER): MISSING, DOES NOT APPLY

CONDITIONING ID:             TMAT0009
DESCRIPTION:                 HOW OFTEN DO STUDENTS USE A COMPUTER (TEACHER)?
GRADES/ASSESSMENTS:          N04, S04, N08, S08
GROUP LABEL:                 T_COMPTR          LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                     T044506           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS             NUMBER OF SPECIFICATION RECORDS:  5
001   T_CMPTR1   (1             )  0000        USE COMPUTER (TEACHER): ALMOST EVERY DAY
002   T_CMPTR2   (2             )  1000        USE COMPUTER (TEACHER): ONCE OR TWICE A WEEK
003   T_CMPTR3   (3             )  0100        USE COMPUTER (TEACHER): ONCE OR TWICE A MONTH
004   T_CMPTR4   (4             )  0010        USE COMPUTER (TEACHER): NEVER OR HARDLY EVER
005   T_CMPTR?   (M,DNA         )  0001        USE COMPUTER (TEACHER): MISSING, DOES NOT APPLY

CONDITIONING ID:             TMAT0010
DESCRIPTION:                 HOW OFTEN DO STUDENTS WRITE ABOUT PROBLEM SOLVING (TEACHER)?
GRADES/ASSESSMENTS:          N04, S04, N08, S08
GROUP LABEL:                 T_PRBSOL          LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                     T044507           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS             NUMBER OF SPECIFICATION RECORDS:  5
```

Table F-3 (continued)
Conditioning Variables for Mathematics Main Samples

```
001   T_PRBSL1    (1           )  0000        PROBLEM SOLVING (TEACHER): ALMOST EVERY DAY
002   T_PRBSL2    (2           )  1000        PROBLEM SOLVING (TEACHER): ONCE OR TWICE A WEEK
003   T_PRBSL3    (3           )  0100        PROBLEM SOLVING (TEACHER): ONCE OR TWICE A MONTH
004   T_PRBSL4    (4           )  0010        PROBLEM SOLVING (TEACHER): NEVER OR HARDLY EVER
005   T_PRBSL?    (M,DNA       )  0001        PROBLEM SOLVING (TEACHER): MISSING, DOES NOT APPLY

CONDITIONING ID:          TMAT0011
DESCRIPTION:              HOW OFTEN DO STUDENTS WRITE REPORTS/DO PROJECTS (TEACHER)?
GRADES/ASSESSMENTS:       N04, S04, N08, S08
GROUP LABEL:              T_REPPRJ           LENGTH OF CONTRAST FIELD      :  4
NAEP ID:                  T044508            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS              NUMBER OF SPECIFICATION RECORDS:  5
001   T_REPPJ1    (1           )  0000       REPORTS/PROJECTS (TEACHER): ALMOST EVERY DAY
002   T_REPPJ2    (2           )  1000       REPORTS/PROJECTS (TEACHER): ONCE OR TWICE A WEEK
003   T_REPPJ3    (3           )  0100       REPORTS/PROJECTS (TEACHER): ONCE OR TWICE A MONTH
004   T_REPPJ4    (4           )  0010       REPORTS/PROJECTS (TEACHER): NEVER OR HARDLY EVER
005   T_REPPJ?    (M,DNA       )  0001       REPORTS/PROJECTS (TEACHER): MISSING, DOES NOT APPLY

CONDITIONING ID:          TMAT0012
DESCRIPTION:              HOW OFTEN DO STUDENTS DISCUSS MATH WITH OTHER STUDENTS (TEACHER)?
GRADES/ASSESSMENTS:       N04, S04, N08, S08
GROUP LABEL:              T_DISMAT           LENGTH OF CONTRAST FIELD      :  4
NAEP ID:                  T044509            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS              NUMBER OF SPECIFICATION RECORDS:  5
001   T_DISMA1    (1           )  0000       STUDENTS DISCUSS MATH (TEACHER): ALMOST EVERY DAY
002   T_DISMA2    (2           )  1000       STUDENTS DISCUSS MATH (TEACHER): ONCE OR TWICE A WEEK
003   T_DISMA3    (3           )  0100       STUDENTS DISCUSS MATH (TEACHER): ONCE OR TWICE A MONTH
004   T_DISMA4    (4           )  0010       STUDENTS DISCUSS MATH (TEACHER): NEVER OR HARDLY EVER
005   T_DISMA?    (M,DNA       )  0001       STUDENTS DISCUSS MATH (TEACHER): MISSING, DOES NOT APPLY

CONDITIONING ID:          TMAT0013
DESCRIPTION:              HOW OFTEN DO STUDENTS WORK REAL LIFE PROBLEMS (TEACHER)?
GRADES/ASSESSMENTS:       N04, S04, N08, S08
GROUP LABEL:              T_RLPROB           LENGTH OF CONTRAST FIELD      :  4
NAEP ID:                  T044510            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS              NUMBER OF SPECIFICATION RECORDS:  5
001   T_RLPRB1    (1           )  0000       REAL LIFE PROBLEMS (TEACHER): ALMOST EVERY DAY
002   T_RLPRB2    (2           )  1000       REAL LIFE PROBLEMS (TEACHER): ONCE OR TWICE A WEEK
003   T_RLPRB3    (3           )  0100       REAL LIFE PROBLEMS (TEACHER): ONCE OR TWICE A MONTH
004   T_RLPRB4    (4           )  0010       REAL LIFE PROBLEMS (TEACHER): NEVER OR HARDLY EVER
005   T_RLPRB?    (M,DNA       )  0001       REAL LIFE PROBLEMS (TEACHER): MISSING, DOES NOT APPLY

CONDITIONING ID:          TMAT0014
DESCRIPTION:              HOW OFTEN DO STUDENTS MAKE UP MATH PROBLEMS (TEACHER)?
GRADES/ASSESSMENTS:       N04, S04, N08, S08
GROUP LABEL:              T_MUPROB           LENGTH OF CONTRAST FIELD      :  4
NAEP ID:                  T044511            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS              NUMBER OF SPECIFICATION RECORDS:  5
001   T_MUPRB1    (1           )  0000       STUDENTS MAKE UP PROBLEMS (TEACHER): ALMOST EVERY DAY
002   T_MUPRB2    (2           )  1000       STUDENTS MAKE UP PROBLEMS (TEACHER): ONCE OR TWICE A WEEK
003   T_MUPRB3    (3           )  0100       STUDENTS MAKE UP PROBLEMS (TEACHER): ONCE OR TWICE A MONTH
004   T_MUPRB4    (4           )  0010       STUDENTS MAKE UP PROBLEMS (TEACHER): NEVER OR HARDLY EVER
005   T_MUPRB?    (M,DNA       )  0001       STUDENTS MAKE UP PROBLEMS (TEACHER): MISSING, DOES NOT APPLY

CONDITIONING ID:          TMAT0015
DESCRIPTION:              HOW MUCH EMPHASIS ON NUMBERS AND OPERATIONS?
GRADES/ASSESSMENTS:       N04, S04, N08, S08
GROUP LABEL:              EMP_N&OP           LENGTH OF CONTRAST FIELD      :  3
NAEP ID:                  T044601            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS              NUMBER OF SPECIFICATION RECORDS:  4
001   EMP_N&O1    (1           )  000        EMPHASIS NUMBERS AND OPERATIONS:  HEAVY EMPHASIS
002   EMP_N&O2    (2           )  100        EMPHASIS NUMBERS AND OPERATIONS:  MODERATE EMPHASIS
003   EMP_N&O3    (3           )  010        EMPHASIS NUMBERS AND OPERATIONS:  LITTLE OR NONE
004   EMP_N&O?    (M,DNA       )  001        EMPHASIS NUMBERS AND OPERATIONS:  MISSING, DOES NOT APPLY
```

Table F-3 (continued)
Conditioning Variables for Mathematics Main Samples

```
CONDITIONING ID:          TMAT0016
DESCRIPTION:              HOW MUCH EMPHASIS ON MEASUREMENT?
GRADES/ASSESSMENTS:       N04, S04, N08, S08
GROUP LABEL:              EMP_MEAS            LENGTH OF CONTRAST FIELD      :  3
NAEP ID:                  T044602             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   EMP_MEA1   (1          )  000           EMPHASIS MEASUREMENT:  HEAVY EMPHASIS
002   EMP_MEA2   (2          )  100           EMPHASIS MEASUREMENT:  MODERATE EMPHASIS
003   EMP_MEA3   (3          )  010           EMPHASIS MEASUREMENT:  LITTLE OR NONE
004   EMP_MEA?   (M,DNA      )  001           EMPHASIS MEASUREMENT:  MISSING, DOES NOT APPLY

CONDITIONING ID:          TMAT0017
DESCRIPTION:              HOW MUCH EMPHASIS ON GEOMETRY?
GRADES/ASSESSMENTS:       N04, S04, N08, S08
GROUP LABEL:              EMP_GEOM            LENGTH OF CONTRAST FIELD      :  3
NAEP ID:                  T044603             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   EMP_GEO1   (1          )  000           EMPHASIS GEOMETRY:  HEAVY EMPHASIS
002   EMP_GEO2   (2          )  100           EMPHASIS GEOMETRY:  MODERATE EMPHASIS
003   EMP_GEO3   (3          )  010           EMPHASIS GEOMETRY:  LITTLE OR NONE
004   EMP_GEO?   (M,DNA      )  001           EMPHASIS GEOMETRY:  MISSING, DOES NOT APPLY

CONDITIONING ID:          TMAT0018
DESCRIPTION:              HOW MUCH EMPHASIS ON DATA ANALYSIS/STATISTICS/PROBABILITY? (GRADE 4)
GRADES/ASSESSMENTS:       N04, S04
GROUP LABEL:              EMP_DSP4            LENGTH OF CONTRAST FIELD      :  3
NAEP ID:                  T044611             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   EMP_DSP1   (1          )  000           EMPHASIS DATA ANALYSIS/STAT/PROB:  HEAVY EMPHASIS
002   EMP_DSP2   (2          )  100           EMPHASIS DATA ANALYSIS/STAT/PROB:  MODERATE EMPHASIS
003   EMP_DSP3   (3          )  010           EMPHASIS DATA ANALYSIS/STAT/PROB:  LITTLE OR NONE
004   EMP_DSP?   (M,DNA      )  001           EMPHASIS DATA ANALYSIS/STAT/PROB:  MISSING, DOES NOT APPLY

CONDITIONING ID:          TMAT0019
DESCRIPTION:              HOW MUCH EMPHASIS ON DATA ANALYSIS/STATISTICS/PROBABILITY? (GRADE 8)
GRADES/ASSESSMENTS:       N08, S08
GROUP LABEL:              EMP_DSP8            LENGTH OF CONTRAST FIELD      :  3
NAEP ID:                  T044604             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   EMP_DSP1   (1          )  000           EMPHASIS DATA ANALYSIS/STAT/PROB:  HEAVY EMPHASIS
002   EMP_DSP2   (2          )  100           EMPHASIS DATA ANALYSIS/STAT/PROB:  MODERATE EMPHASIS
003   EMP_DSP3   (3          )  010           EMPHASIS DATA ANALYSIS/STAT/PROB:  LITTLE OR NONE
004   EMP_DSP?   (M,DNA      )  001           EMPHASIS DATA ANALYSIS/STAT/PROB:  MISSING, DOES NOT APPLY

CONDITIONING ID:          TMAT0020
DESCRIPTION:              HOW MUCH EMPHASIS ON ALGEBRA AND FUNCTIONS? (GRADE 4)
GRADES/ASSESSMENTS:       N04, S04
GROUP LABEL:              EMP_ALG4            LENGTH OF CONTRAST FIELD      :  3
NAEP ID:                  T044612             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   EMP_ALG1   (1          )  000           EMPHASIS ALGEBRA AND FUNCTIONS:  HEAVY EMPHASIS
002   EMP_ALG2   (2          )  100           EMPHASIS ALGEBRA AND FUNCTIONS:  MODERATE EMPHASIS
003   EMP_ALG3   (3          )  010           EMPHASIS ALGEBRA AND FUNCTIONS:  LITTLE OR NONE
004   EMP_ALG?   (M,DNA      )  001           EMPHASIS ALGEBRA AND FUNCTIONS:  MISSING, DOES NOT APPLY

CONDITIONING ID:          TMAT0021
DESCRIPTION:              HOW MUCH EMPHASIS ON ALGEBRA AND FUNCTIONS? (GRADE 8)
GRADES/ASSESSMENTS:       N08, S08
GROUP LABEL:              EMP_ALG8            LENGTH OF CONTRAST FIELD      :  3
NAEP ID:                  T044605             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   EMP_ALG1   (1          )  000           EMPHASIS ALGEBRA AND FUNCTIONS:  HEAVY EMPHASIS
002   EMP_ALG2   (2          )  100           EMPHASIS ALGEBRA AND FUNCTIONS:  MODERATE EMPHASIS
003   EMP_ALG3   (3          )  010           EMPHASIS ALGEBRA AND FUNCTIONS:  LITTLE OR NONE
004   EMP_ALG?   (M,DNA      )  001           EMPHASIS ALGEBRA AND FUNCTIONS:  MISSING, DOES NOT APPLY
```

BEST COPY AVAILABLE

Table F-3 (continued)
Conditioning Variables for Mathematics Main Samples

```
CONDITIONING ID:          TMAT0022
DESCRIPTION:              HOW MUCH EMPHASIS ON LEARNING FACTS/CONCEPTS?
GRADES/ASSESSMENTS:       N04, S04, N08, S08
GROUP LABEL:              EMP_F/C             LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                  T044606             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   EMP_F/C1   (1          )  000           EMPHASIS FACTS/CONCEPTS:  HEAVY EMPHASIS
002   EMP_F/C2   (2          )  100           EMPHASIS FACTS/CONCEPTS:  MODERATE EMPHASIS
003   EMP_F/C3   (3          )  010           EMPHASIS FACTS/CONCEPTS:  LITTLE OR NONE
004   EMP_F/C?   (M,DNA      )  001           EMPHASIS FACTS/CONCEPTS:  MISSING, DOES NOT APPLY

CONDITIONING ID:          TMAT0023
DESCRIPTION:              HOW MUCH EMPHASIS ON LEARNING SKILLS/PROCEDURES?
GRADES/ASSESSMENTS:       N04, S04, N08, S08
GROUP LABEL:              EMP_S/P             LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                  T044607             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   EMP_S/P1   (1          )  000           EMPHASIS SKILLS/PROCEDURES: HEAVY EMPHASIS
002   EMP_S/P2   (2          )  100           EMPHASIS SKILLS/PROCEDURES: MODERATE EMPHASIS
003   EMP_S/P3   (3          )  010           EMPHASIS SKILLS/PROCEDURES: LITTLE OR NONE
004   EMP_S/P?   (M,DNA      )  001           EMPHASIS SKILLS/PROCEDURES: MISSING, DOES NOT APPLY

CONDITIONING ID:          TMAT0024
DESCRIPTION:              HOW MUCH EMPHASIS ON REASONING/ANALYSIS?
GRADES/ASSESSMENTS:       N04, S04, N08, S08
GROUP LABEL:              EMP_R/A             LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                  T044608             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   EMP_R/A1   (1          )  000           EMPHASIS REASONING/ANALYSIS: HEAVY EMPHASIS
002   EMP_R/A2   (2          )  100           EMPHASIS REASONING/ANALYSIS: MODERATE EMPHASIS
003   EMP_R/A3   (3          )  010           EMPHASIS REASONING/ANALYSIS: LITTLE OR NONE
004   EMP_R/A?   (M,DNA      )  001           EMPHASIS REASONING/ANALYSIS: MISSING, DOES NOT APPLY

CONDITIONING ID:          TMAT0025
DESCRIPTION:              HOW MUCH EMPHASIS ON COMMUNICATING MATH IDEAS?
GRADES/ASSESSMENTS:       N04, S04, N08, S08
GROUP LABEL:              EMP_CMI             LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                  T044609             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   EMP_CMI1   (1          )  000           EMPHASIS COMMUNICATING MATH IDEAS: HEAVY EMPHASIS
002   EMP_CMI2   (2          )  100           EMPHASIS COMMUNICATING MATH IDEAS: MODERATE EMPHASIS
003   EMP_CMI3   (3          )  010           EMPHASIS COMMUNICATING MATH IDEAS: LITTLE OR NONE
004   EMP_CMI?   (M,DNA      )  001           EMPHASIS COMMUNICATING MATH IDEAS: MISSING, DOES NOT APPLY

CONDITIONING ID:          TMAT0026
DESCRIPTION:              HOW MUCH EMPHASIS ON APPRECIATING MATHEMATICS?
GRADES/ASSESSMENTS:       N04, S04, N08, S08
GROUP LABEL:              EMP_MAP             LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                  T044610             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   EMP_MAP1   (1          )  000           EMPHASIS MATH APPRECIATION: HEAVY EMPHASIS
002   EMP_MAP2   (2          )  100           EMPHASIS MATH APPRECIATION: MODERATE EMPHASIS
003   EMP_MAP3   (3          )  010           EMPHASIS MATH APPRECIATION: LITTLE OR NONE
004   EMP_MAP?   (M,DNA      )  001           EMPHASIS MATH APPRECIATION: MISSING, DOES NOT APPLY

CONDITIONING ID:          TMAT0027
DESCRIPTION:              DO YOU PERMIT UNRESTRICTED USE OF CALCULATORS? (TEACHER)
GRADES/ASSESSMENTS:       N04, S04, N08, S08
GROUP LABEL:              CALC_UNR            LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                  T045401             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS               NUMBER OF SPECIFICATION RECORDS:  3
001   CALUNR-Y   (1          )  00            UNRESTRICTED CALCULATOR USE:  YES
002   CALUNR-N   (2          )  10            UNRESTRICTED CALCULATOR USE:  NO
003   CALUNR-?   (M,DNA      )  01            UNRESTRICTED CALCULATOR USE:  MISSING, DOES NOT APPLY
```

Table F-3 (continued)
Conditioning Variables for Mathematics Main Samples

```
CONDITIONING ID:          TMAT0028
DESCRIPTION:              DO YOU PERMIT USE OF CALCULATORS ON TESTS? (TEACHER)
GRADES/ASSESSMENTS:       N04, S04, N08, S08
GROUP LABEL:              CALC_TST            LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                  T044801             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS               NUMBER OF SPECIFICATION RECORDS:  3
001   CALTST-Y   (1          )  00            CALCULATOR USE ON TESTS:  YES
002   CALTST-N   (2          )  10            CALCULATOR USE ON TESTS:  NO
003   CALTST-?   (M,DNA      )  01            CALCULATOR USE ON TESTS:  MISSING, DOES NOT APPLY

CONDITIONING ID:          TMAT0029
DESCRIPTION:              TEACHER HOURS SPENT IN IN-SERVICE MATHEMATICS EDUCATION
GRADES/ASSESSMENTS:       N04, S04, N08, S08
GROUP LABEL:              INSERV_M            LENGTH OF CONTRAST FIELD       :  5
NAEP ID:                  T040901             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS               NUMBER OF SPECIFICATION RECORDS:  6
001   INSERVM1   (1          )  00000         MATH HOURS IN-SERVICE:  NONE
002   INSERVM2   (2          )  10000         MATH HOURS IN-SERVICE:  LESS THAN 6 HOURS
003   INSERVM3   (3          )  01000         MATH HOURS IN-SERVICE:  6-15 HOURS
004   INSERVM4   (4          )  00100         MATH HOURS IN-SERVICE:  16-35 HOURS
005   INSERVM5   (5          )  00010         MATH HOURS IN-SERVICE:  MORE THAN 35 HOURS
006   INSERVM?   (M,DNA      )  00001         MATH HOURS IN-SERVICE:  MISSING, DOES NOT APPLY

CONDITIONING ID:          TMAT0030
DESCRIPTION:              ONE OR MORE COLLEGE COURSES IN SEVEN SUBJECTS
GRADES/ASSESSMENTS:       N04, S04, N08, S08
GROUP LABEL:              T_MATHCR            LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                  TMATHCR             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS               NUMBER OF SPECIFICATION RECORDS:  3
001   T_MATCR1   (1,M        )  00            TEACHER COURSES:  0-3
002   T_MATCR2   (2          )  10            TEACHER COURSES:  4-5
003   T_MATCR3   (3          )  01            TEACHER COURSES:  6-7

CONDITIONING ID:          TMAT0031
DESCRIPTION:              COLLEGE OR IN-SERVICE TRAINING IN SEVEN AREAS
GRADES/ASSESSMENTS:       N04, S04, N08, S08
GROUP LABEL:              T_MATHTR            LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                  TMATHTR             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS               NUMBER OF SPECIFICATION RECORDS:  3
001   T_MATTR1   (1,M        )  00            TEACHER TRAINING:  0-3
002   T_MATTR2   (2          )  10            TEACHER TRAINING:  4-5
003   T_MATTR3   (3          )  01            TEACHER TRAINING:  6-7
```

Table F-4
Conditioning Variables for Writing Main Samples

```
CONDITIONING ID:          BACK0001
DESCRIPTION:              GRAND MEAN
GRADES/ASSESSMENTS:       N04, S04, N08, S08, N12
GROUP LABEL:              OVERALL            LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                  BKSER              DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         SCALE              NUMBER OF SPECIFICATION RECORDS:  1
001   OVERALL    (@          )  1            GRAND MEAN

CONDITIONING ID:          BACK0002
DESCRIPTION:              GENDER (DERIVED)
GRADES/ASSESSMENTS:       N04, S04, N08, S08, N12
GROUP LABEL:              GENDER             LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                  DSEX               DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS              NUMBER OF SPECIFICATION RECORDS:  2
001   MALE       (1          )  0            GENDER:  MALE
002   FEMALE     (2          )  1            GENDER:  FEMALE

CONDITIONING ID:          BACK0003
DESCRIPTION:              ETHNICITY/RACE (DERIVED)
GRADES/ASSESSMENTS:       N04, S04, N08, S08, N12
GROUP LABEL:              ETHNICTY           LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                  DRACE              DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS              NUMBER OF SPECIFICATION RECORDS:  4
001   WHIT/AOM   (1,5,6,M    )  000          ETHNICITY:  WHITE, AMERICAN INDIAN, UNCLASSIFIED, MISSING
002   BLACK      (2          )  100          ETHNICITY:  BLACK
003   HISPANIC   (3          )  010          ETHNICITY:  HISPANIC
004   ASIAN      (4          )  001          ETHNICITY:  ASIAN AMERICAN

CONDITIONING ID:          BACK0004
DESCRIPTION:              SIZE AND TYPE OF COMMUNITY
GRADES/ASSESSMENTS:       N04, N08, N12
GROUP LABEL:              STOC               LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                  STOC               DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS              NUMBER OF SPECIFICATION RECORDS:  3
001   LO_METRO   (2          )  00           STOC:  LOW METROPOLITAN
002   HI_METRO   (3          )  10           STOC:  HIGH METROPOLITAN
003   STOC-OTH   (1,4-7,M    )  01           STOC:  OTHER

CONDITIONING ID:          BACK0006
DESCRIPTION:              REGION OF THE COUNTRY
GRADES/ASSESSMENTS:       N04, N08, N12
GROUP LABEL:              REGION             LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                  REGION             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS              NUMBER OF SPECIFICATION RECORDS:  4
001   N_EAST     (1          )  000          REGION:  NORTHEAST
002   S_EAST     (2          )  100          REGION:  SOUTHEAST
003   CENTRAL    (3          )  010          REGION:  CENTRAL
004   WEST       (4          )  001          REGION:  WEST

CONDITIONING ID:          BACK0007
DESCRIPTION:              PARENTS' HIGHEST LEVEL OF EDUCATION
GRADES/ASSESSMENTS:       N04, S04, N08, S08, N12
GROUP LABEL:              PARED              LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                  PARED              DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS              NUMBER OF SPECIFICATION RECORDS:  5
001   <HI_SCH    (1          )  0000         PARED:  LESS THAN HIGH SCHOOL
002   HS_GRAD    (2          )  1000         PARED:  HIGH SCHOOL GRADUATE
003   POST_HS    (3          )  0100         PARED:  POST HIGH SCHOOL
004   COL_GRAD   (4          )  0010         PARED:  COLLEGE GRADUATE
005   PARED-?    (M,IDK      )  0001         PARED:  MISSING, I DON'T KNOW

CONDITIONING ID:          BACK0008
DESCRIPTION:              ITEMS IN THE HOME (NEWSPAPER, > 25 BOOKS, ENCYCLOPEDIA, MAGAZINES)
GRADES/ASSESSMENTS:       N04, S04, N08, S08, N12
```

Table F-4 (continued)
Conditioning Variables for Writing Main Samples

```
GROUP LABEL:                HOMEITMS          LENGTH OF CONTRAST FIELD       :   2
NAEP ID:                    HOMEEN2           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS             NUMBER OF SPECIFICATION RECORDS:   3
001   HITEM<=2   (1,M         )  00           ITEMS IN HOME:  ZERO TO TWO ITEMS, MISSING
002   HITEM=3    (2           )  10           ITEMS IN HOME:  THREE ITEMS
003   HITEM=4    (3           )  01           ITEMS IN HOME:  FOUR ITEMS

CONDITIONING ID:            BACK0009
DESCRIPTION:                HOURS OF TV WATCHING (LINEAR)
GRADES/ASSESSMENTS:         N04, S04, N08, S08, N12
GROUP LABEL:                TVWATCHL          LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                    B001801           DEGREES OF FREEDOM PER CONTRAST:   6
TYPE OF CONTRAST:           SCALE             NUMBER OF SPECIFICATION RECORDS:   7
001   TV-LIN1    (1           )  0            TV WATCHING (LINEAR):  NONE
002   TV-LIN2    (2           )  1            TV WATCHING (LINEAR):  ONE HOUR OR LESS PER DAY
003   TV-LIN3    (3           )  2            TV WATCHING (LINEAR):  TWO HOURS PER DAY
004   TV-LIN4    (4,M         )  3            TV WATCHING (LINEAR):  THREE HOURS PER DAY
005   TV-LIN5    (5           )  4            TV WATCHING (LINEAR):  FOUR HOURS PER DAY
006   TV-LIN6    (6           )  5            TV WATCHING (LINEAR):  FIVE HOURS PER DAY
007   TV-LIN7    (7           )  6            TV WATCHING (LINEAR):  SIX OR MORE HOURS PER DAY

CONDITIONING ID:            BACK0010
DESCRIPTION:                HOURS OF TV WATCHING (QUADRATIC)
GRADES/ASSESSMENTS:         N04, S04, N08, S08, N12
GROUP LABEL:                TVWATCHQ          LENGTH OF CONTRAST FIELD       :   2
NAEP ID:                    B001801           DEGREES OF FREEDOM PER CONTRAST:   6
TYPE OF CONTRAST:           SCALE             NUMBER OF SPECIFICATION RECORDS:   7
001   TV-QUAD1   (1           )  00           TV WATCHING (QUADRATIC):  NONE
002   TV-QUAD2   (2           )  01           TV WATCHING (QUADRATIC):  ONE HOUR OR LESS PER DAY
003   TV-QUAD3   (3           )  04           TV WATCHING (QUADRATIC):  TWO HOURS PER DAY
004   TV-QUAD4   (4,M         )  09           TV WATCHING (QUADRATIC):  THREE HOURS PER DAY
005   TV-QUAD5   (5           )  16           TV WATCHING (QUADRATIC):  FOUR HOURS PER DAY
006   TV-QUAD6   (6           )  25           TV WATCHING (QUADRATIC):  FIVE HOURS PER DAY
007   TV-QUAD7   (7           )  36           TV WATCHING (QUADRATIC):  SIX OR MORE HOURS PER DAY

CONDITIONING ID:            BACK0011
DESCRIPTION:                HOME LANGUAGE MINORITY (HOW OFTEN PEOPLE IN HOME SPEAK OTHER THAN ENGLISH?)
GRADES/ASSESSMENTS:         N04, S04, N08, S08, N12
GROUP LABEL:                HOMELANG          LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                    B003201           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS             NUMBER OF SPECIFICATION RECORDS:   2
001   HL-NEV/?   (1,M         )  0            HOME LANGUAGE MINORITY:  NEVER, MISSING
002   HL-SM/AL   (2,3         )  1            HOME LANGUAGE MINORITY:  SOMTIMES, ALWAYS

CONDITIONING ID:            BACK0012
DESCRIPTION:                HOMEWORK ASSIGNED? (GRADE 4)
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                HW-CORE4          LENGTH OF CONTRAST FIELD       :   2
NAEP ID:                    B006601           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS             NUMBER OF SPECIFICATION RECORDS:   3
001   HW4-MISS   (M           )  00           HOMEWORK ASSIGNED?:  MISSING
002   HW4-NONE   (1           )  10           HOMEWORK ASSIGNED?:  NO HOMEWORK ASSIGNED
003   HW4-YES    (2-5         )  01           HOMEWORK ASSIGNED?:  YES

CONDITIONING ID:            BACK0013
DESCRIPTION:                AMOUNT OF HOMEWORK (LINEAR) (GRADE 4)
GRADES/ASSESSMENTS:         N04, S04
GROUP LABEL:                HMWRKL4           LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                    B006601           DEGREES OF FREEDOM PER CONTRAST:   3
TYPE OF CONTRAST:           SCALE             NUMBER OF SPECIFICATION RECORDS:   4
001   HW4-LIN1   (1,2,M       )  0            AMOUNT OF HOMEWORK (LINEAR):  DON'T HAVE, DON'T DO, MISSING
002   HW4-LIN2   (3           )  1            AMOUNT OF HOMEWORK (LINEAR):  ONE HALF HOUR
003   HW4-LIN3   (4           )  2            AMOUNT OF HOMEWORK (LINEAR):  ONE HOUR
004   HW4-LIN4   (5           )  3            AMOUNT OF HOMEWORK (LINEAR):  MORE THAN ONE HOUR
```

Table F-4 (continued)
Conditioning Variables for Writing Main Samples

```
CONDITIONING ID:           BACK0014
DESCRIPTION:               AMOUNT OF HOMEWORK (QUADRATIC) (GRADE 4)
GRADES/ASSESSMENTS:        N04, S04
GROUP LABEL:               HMWRKQ4              LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                   B006601              DEGREES OF FREEDOM PER CONTRAST:  3
TYPE OF CONTRAST:          SCALE                NUMBER OF SPECIFICATION RECORDS:  4
001   HW4QUAD1   (1,2,M       )  0              AMOUNT OF HOMEWORK (QUADRATIC):  DON'T HAVE, DON'T DO,MISS
002   HW4QUAD2   (3           )  1              AMOUNT OF HOMEWORK (QUADRATIC):  ONE HALF HOUR
003   HW4QUAD3   (4           )  4              AMOUNT OF HOMEWORK (QUADRATIC):  ONE HOUR
004   HW4QUAD4   (5           )  9              AMOUNT OF HOMEWORK (QUADRATIC):  MORE THAN ONE HOUR

CONDITIONING ID:           BACK0015
DESCRIPTION:               HOMEWORK ASSIGNED?
GRADES/ASSESSMENTS:        N08, S08, N12
GROUP LABEL:               HW-CORE              LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                   B003901              DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS                NUMBER OF SPECIFICATION RECORDS:  3
001   HWC-MISS   (M           )  00             HOMEWORK ASSIGNED?:  MISSING
002   HWC-NONE   (1           )  10             HOMEWORK ASSIGNED?:  NO HOMEWORK ASSIGNED
003   HWC-YES    (2-6         )  01             HONEWORK ASSIGNED?:  YES

CONDITIONING ID:           BACK0016
DESCRIPTION:               AMOUNT OF HOMEWORK (LINEAR)
GRADES/ASSESSMENTS:        N08, S08, N12
GROUP LABEL:               HMWRKL               LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                   B003901              DEGREES OF FREEDOM PER CONTRAST:  4
TYPE OF CONTRAST:          SCALE                NUMBER OF SPECIFICATION RECORDS:  5
001   HW-LIN1    (1,2,M       )  0              AMOUNT OF HOMEWORK (LINEAR):  DON'T HAVE, DON'T DO,MISS
002   HW-LIN2    (3           )  1              AMOUNT OF HOMEWORK (LINEAR):  ONE HALF HOUR
003   HW-LIN3    (4           )  2              AMOUNT OF HOMEWORK (LINEAR):  ONE HOUR
004   HW-LIN4    (5           )  3              AMOUNT OF HOMEWORK (LINEAR):  TWO HOURS
005   HW-LIN5    (6           )  4              AMOUNT OF HOMEWORK (LINEAR):  MORE THAN TWO HOURS

CONDITIONING ID:           BACK0017
DESCRIPTION:               AMOUNT OF HOMEWORK (QUADRATIC)
GRADES/ASSESSMENTS:        N08, S08, N12
GROUP LABEL:               HMWRKQ               LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                   B003901              DEGREES OF FREEDOM PER CONTRAST:  4
TYPE OF CONTRAST:          SCALE                NUMBER OF SPECIFICATION RECORDS:  5
001   HW-QUAD1   (1,2,M       )  00             AMOUNT OF HOMEWORK (QUADRATIC):  DON'T HAVE, DON'T DO,MISS
002   HW-QUAD2   (3           )  01             AMOUNT OF HOMEWORK (QUADRATIC):  ONE HALF HOUR
003   HW-QUAD3   (4           )  04             AMOUNT OF HOMEWORK (QUADRATIC):  ONE HOUR
004   HW-QUAD4   (5           )  09             AMOUNT OF HOMEWORK (QUADRATIC):  TWO HOURS
005   HW-QUAD5   (6           )  16             AMOUNT OF HOMEWORK (QUADRATIC):  MORE THAN TWO HOURS

CONDITIONING ID:           BACK0018
DESCRIPTION:               PERCENT WHITE STUDENTS IN SCHOOL
GRADES/ASSESSMENTS:        N04, S04, N08, S08, N12
GROUP LABEL:               %WHITE               LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                   PCTWHT               DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS                NUMBER OF SPECIFICATION RECORDS:  3
001   PREDOM/?   (80-110,M    )  00             PREDOMINANTLY WHITE, MISSING
002   MINORITY   (0-49        )  10             WHITE MINORITY
003   INTEGRAT   (50-79       )  01             INTEGRATED

CONDITIONING ID:           BACK0019
DESCRIPTION:               MODAL AGE, MODAL GRADE (DERIVED)
GRADES/ASSESSMENTS:        N04, N08, N12
GROUP LABEL:               AGE/GRAD             LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                   MODGRAG              DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS                NUMBER OF SPECIFICATION RECORDS:  5
001   <MA/=MG    (1           )  0000           LESS THAN MODAL AGE, MODAL GRADE
002   =MA/<MG    (2           )  1000           MODAL AGE, LESS THAN MODAL GRADE
003   =MA/=MG    (3,M         )  0100           MODAL AGE, MODAL GRADE, MISSING
004   =MA/>MG    (4           )  0010           MODAL AGE, GREATER THAN MODAL GRADE
```

Table F-4 (continued)
Conditioning Variables for Writing Main Samples

```
005   >MA/=MG    (5           )  0001          GREATER THAN MODAL AGE, MODAL GRADE

CONDITIONING ID:           BACK0020
DESCRIPTION:               SCHOOL TYPE:  PUBLIC/NON-PUBLIC
GRADES/ASSESSMENTS:        N04, N08, N12
GROUP LABEL:               SCH_TYPE            LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                   SCHTYPE             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS               NUMBER OF SPECIFICATION RECORDS:   2
001   PUBLIC     (1           )  0             PUBLIC SCHOOL
002   NON_PUBL   (2-5,M       )  1             PRIVATE, CATHOLIC, BIA, DEPARTMENT OF DEFENSE, MISSING

CONDITIONING ID:           BACK0021
DESCRIPTION:               SINGLE/MULTIPLE PARENT(S) AT HOME
GRADES/ASSESSMENTS:        N04, S04, N08, S08, N12
GROUP LABEL:               PARENTS             LENGTH OF CONTRAST FIELD       :   2
NAEP ID:                   SINGLEP             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS               NUMBER OF SPECIFICATION RECORDS:   3
001   NOT2PARS   (3-4,M       )  00            NOT TWO PARENTS, MISSING
002   2PARENTS   (2           )  10            EITHER FATHER OR MOTHER AT HOME, BUT NOT BOTH
003   2PARENTS   (1           )  01            BOTH FATHER AND MOTHER AT HOME

CONDITIONING ID:           BACK0022
DESCRIPTION:               MOTHER AT HOME
GRADES/ASSESSMENTS:        N04, S04, N08, S08, N12
GROUP LABEL:               MOM@HOME            LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                   B005601             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS               NUMBER OF SPECIFICATION RECORDS:   2
001   MOM@HM-N   (2,M         )  0             MOTHER AT HOME:  NO, MISSING
002   MOM@HM-Y   (1           )  1             MOTHER AT HOME:  YES

CONDITIONING ID:           BACK0023
DESCRIPTION:               PAGES READ FOR SCHOOL AND HOMEWORK EACH DAY (CONTRAST 1)
GRADES/ASSESSMENTS:        N04, S04, N08, S08, N12
GROUP LABEL:               PGSREADW            LENGTH OF CONTRAST FIELD       :   2
NAEP ID:                   B001101             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS               NUMBER OF SPECIFICATION RECORDS:   3
001   <=5_PGS    (5,M         )  00            PAGES READ (W):  5 OR FEWER PAGES, MISSING
002   6-10PGS    (4           )  10            PAGES READ (W):  6-10 PAGES
003   >10_PGS    (1-3         )  01            PAGES READ (W):  MORE THAN 20, 16-20 PAGES, 11-15 PAGES

CONDITIONING ID:           BACK0025
DESCRIPTION:               WENT TO PRESCHOOL?
GRADES/ASSESSMENTS:        N04, S04
GROUP LABEL:               PRESCH              LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                   B004201             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS               NUMBER OF SPECIFICATION RECORDS:   2
001   PRESCH-N   (2,3,IDK,M   )  0             WENT TO PRESCHOOL?:  NO, I DON'T KNOW, MISSING
002   PRESCH-Y   (1           )  1             WENT TO PRESCHOOL?:  YES

CONDITIONING ID:           BACK0026
DESCRIPTION:               DAYS OF SCHOOL MISSED LAST MONTH
GRADES/ASSESSMENTS:        N08, S08, N12
GROUP LABEL:               SCH_MISS            LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                   S004001             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS               NUMBER OF SPECIFICATION RECORDS:   2
001   >=3_DAYS   (3-5,M       )  0             DAYS SCHOOL MISSED:  3 OR 4, 5 TO 10, MORE THAN 10,MISS
002   <=2_DAYS   (1,2         )  1             DAYS SCHOOL MISSED:  NONE, 1 OR 2 DAYS

CONDITIONING ID:           BACK0027
DESCRIPTION:               HIGH SCHOOL PROGRAM GRADES/ASSESSMENTS:        N12
GROUP LABEL:               HS_PROG             LENGTH OF CONTRAST FIELD       :   2
NAEP ID:                   B005001             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS               NUMBER OF SPECIFICATION RECORDS:   3
001   GENERL/?   (1,M         )  00            HS PROGRAM:  GENERAL, MISSING
002   COL_PREP   (2           )  10            HS PROGRAM:  COLLEGE PREPARATORY
```

Table F-4 (continued)
Conditioning Variables for Writing Main Samples

```
003   VOC/TECH   (3           )  01          HS PROGRAM:  VOCATIONAL, TECHNICAL

CONDITIONING ID:              BACK0028
DESCRIPTION:                  NUMBER OF SEMESTERS ENGLISH/LITERATURE/WRITING (MISSING/NOT-MISSING)
GRADES/ASSESSMENTS:           N12
GROUP LABEL:                  #SEM_ENG           LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                      B007101            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:             CLASS              NUMBER OF SPECIFICATION RECORDS:  2
001   #ENG-?     (M           )  0           NUMBER OF SEMESTERS ENGLISH MISSING
002   #ENG-Y     (1-9         )  1           NUMBER OF SEMESTERS ENGLISH NOT-MISSING

CONDITIONING ID:              BACK0029
DESCRIPTION:                  NUMBER OF SEMESTERS ENGLISH/LITERATURE/WRITING (LINEAR)
GRADES/ASSESSMENTS:           N12
GROUP LABEL:                  #ENG-LIN           LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                      B007101            DEGREES OF FREEDOM PER CONTRAST:  8
TYPE OF CONTRAST:             SCALE              NUMBER OF SPECIFICATION RECORDS: 10
001   #ENG-L01   (1           )  1           NUMBER OF SEMESTERS ENGLISH LINEAR
002   #ENG-L02   (2           )  2           NUMBER OF SEMESTERS ENGLISH LINEAR
003   #ENG-L03   (3           )  3           NUMBER OF SEMESTERS ENGLISH LINEAR
004   #ENG-L04   (4           )  4           NUMBER OF SEMESTERS ENGLISH LINEAR
005   #ENG-L05   (5           )  5           NUMBER OF SEMESTERS ENGLISH LINEAR
006   #ENG-L06   (6           )  6           NUMBER OF SEMESTERS ENGLISH LINEAR
007   #ENG-L07   (7           )  7           NUMBER OF SEMESTERS ENGLISH LINEAR
008   #ENG-L08   (8           )  8           NUMBER OF SEMESTERS ENGLISH LINEAR
009   #ENG-L09   (9           )  9           NUMBER OF SEMESTERS ENGLISH LINEAR
010   #ENG-L10   (M           )  0           NUMBER OF SEMESTERS ENGLISH LINEAR (MISSING)

CONDITIONING ID:              BACK0042
DESCRIPTION:                  BORN IN ONE OF THE 50 STATES
GRADES/ASSESSMENTS:           N04, S04, N08, S08, N12
GROUP LABEL:                  BORN_USA           LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                      B007801            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:             CLASS              NUMBER OF SPECIFICATION RECORDS:  2
001   USA-YES    (1           )  0           BORN IN THE USA:  YES
002   USA-NO/?   (2,M         )  1           BORN IN THE USA:  NO/MIS SING

CONDITIONING ID:              BACK0043
DESCRIPTION:                  HOW MANY TIMES CHANGED SCHOOLS IN THE LAST TWO YEARS?
GRADES/ASSESSMENTS:           N04, S04, N08, S08
GROUP LABEL:                  SCH_CHGS           LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                      B007301            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:             CLASS              NUMBER OF SPECIFICATION RECORDS:  4
001   CHGSCH=0   (1           )  000         CHANGED SCHOOLS (NONE)
002   CHGSCH=1   (2           )  100         CHANGED SCHOOLS ONCE
003   CHGSCH=2   (3           )  010         CHANGED SCHOOLS TWICE
004   CHGSCH3+   (4,M         )  001         CHANGED SCHOOLS 3 OR MORE TIMES, MISSING

CONDITIONING ID:              BACK0044
DESCRIPTION:                  HOW MANY GRADES HAVE YOU GONE TO SCHOOL IN THIS STATE? (K-4)
GRADES/ASSESSMENTS:           N04, S04
GROUP LABEL:                  GRDS_ST4           LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                      B007601            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:             CLASS              NUMBER OF SPECIFICATION RECORDS:  3
001   ST4GRD<1   (1,M         )  00          LESS THAN ONE GRADE IN THIS STATE, MISSING (K-4)
002   ST4GRD12   (2           )  10          ONE TO TWO GRADES IN THIS STATE (K-4)
003   ST4GRD3+   (3           )  01          THREE OR MORE GRADES IN THIS STATE (K-4)

CONDITIONING ID:              BACK0045
DESCRIPTION:                  HOW OFTEN DO YOU DISCUSS THINGS STUDIED IN SCHOOL WITH SOMEONE AT HOME?

GRADES/ASSESSMENTS:           N04, S04, N08, S08, N12
GROUP LABEL:                  DISC@HOM           LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                      B007401            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:             CLASS              NUMBER OF SPECIFICATION RECORDS:  4
```

Table F-4 (continued)
Conditioning Variables for Writing Main Samples

```
001   DIS@HOM1   (1             )  000         DISCUSS AT HOME (ALMOST EVERYDAY)
002   DIS@HOM2   (2             )  100         DISCUSS AT HOME (ONCE OR TWICE A WEEK)
003   DIS@HOM3   (3             )  010         DISCUSS AT HOME (ONCE OR TWICE A MONTH)
004   DIS@HOM4   (4,M           )  001         DISCUSS AT HOME (NEVER OR HARDLY EVER, MISSING)

CONDITIONING ID:              BACK0046
DESCRIPTION:                  HOW OFTEN DO USE A COMPUTER FOR SCHOOLWORK?
GRADES/ASSESSMENTS:           N04, S04, N08, S08, N12
GROUP LABEL:                  COMP4SCH         LENGTH OF CONTRAST FIELD      :  4
NAEP ID:                      B007501          DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:             CLASS            NUMBER OF SPECIFICATION RECORDS:  5
001   CMP4SCH1   (1             )  0000        COMPUTER FOR SCHOOLWORK (ALMOST EVERYDAY)
002   CMP4SCH2   (2             )  1000        COMPUTER FOR SCHOOLWORK (ONCE OR TWICE A WEEK)
003   CMP4SCH3   (3             )  0100        COMPUTER FOR SCHOOLWORK (ONCE OR TWICE A MONTH)
004   CMP4SCH4   (4             )  0010        COMPUTER FOR SCHOOLWORK (NEVER OR HARDLY EVER)
005   CMP4SCH5   (M             )  0001        COMPUTER FOR SCHOOLWORK (MISSING)

CONDITIONING ID:              BACK0047
DESCRIPTION:                  HOW MANY GRADES HAVE YOU GONE TO SCHOOL IN THIS STATE? (K-8)
GRADES/ASSESSMENTS:           N08, S08
GROUP LABEL:                  GRDS_ST8         LENGTH OF CONTRAST FIELD      :  3
NAEP ID:                      B007701          DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:             CLASS            NUMBER OF SPECIFICATION RECORDS:  4
001   ST8GRD<1   (1,M           )  000         LESS THAN ONE GRADE IN THIS STATE, MISSING (K-8)
002   ST8GRD12   (2             )  100         ONE TO TWO GRADES IN THIS STATE (K-8)
003   ST8GRD35   (3             )  010         THREE TO FIVE GRADES IN THIS STATE (K-8)
004   ST8GRD>5   (4             )  001         MORE THAN FIVE GRADES IN THIS STATE (K-8)

CONDITIONING ID:              BACK0048
DESCRIPTION:                  WHAT DO YOU EXPECT YOUR MAIN ACTIVITY WILL BE THE YEAR AFTER SCHOOL?
GRADES/ASSESSMENTS:           N12
GROUP LABEL:                  MAINACT          LENGTH OF CONTRAST FIELD      :  5
NAEP ID:                      B007201          DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:             CLASS            NUMBER OF SPECIFICATION RECORDS:  6
001   WORK_F/T   (1             )  00000       MAIN ACTIVITY:  WORK FULL-TIME
002   VOC/BUSI   (2             )  10000       MAIN ACTIVITY:  ATTEND VOC, TECH, BUSINESS SCHOOL
003   2-YR_COL   (3             )  01000       MAIN ACTIVITY:  2-YEAR COLLEGE
004   4-YR_COL   (4             )  00100       MAIN ACTIVITY:  4-YEAR COLLEGE
005   MILITARY   (5             )  00010       MAIN ACTIVITY:  MILITARY
006   OTHERACT   (6,M           )  00001       MAIN ACTIVITY:  OTHER, MISSING

CONDITIONING ID:              BACK0049
DESCRIPTION:                  LEARNING DISRUPTED BY STUDENTS?
GRADES/ASSESSMENTS:           N08, N12
GROUP LABEL:                  LRN_DISR         LENGTH OF CONTRAST FIELD      :  5
NAEP ID:                      B007901          DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:             CLASS            NUMBER OF SPECIFICATION RECORDS:  6
001   LRN_S_AG   (1             )  00000       LEARNING DISRUPTED: STRONGLY AGREE
002   LRN__AG    (2             )  10000       LEARNING DISRUPTED: AGREE
003   LRN__UND   (3             )  01000       LEARNING DISRUPTED: UNDECIDED
004   LRN__DAG   (4             )  00100       LEARNING DISRUPTED: DISAGREE
005   LRN_SDAG   (5             )  00010       LEARNING DISRUPTED: STRONGLY DISAGREE
006   LRN_MISS   (M             )  00001       LEARNING DISRUPTED: MISSING, OMIT

CONDITIONING ID:              BACK0050
DESCRIPTION:                  TEACHERS INTERESTED IN STUDENTS?
GRADES/ASSESSMENTS:           N08, N12
GROUP LABEL:                  TCH INTR         LENGTH OF CONTRAST FIELD      :  5
NAEP ID:                      B007902          DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:             CLASS            NUMBER OF SPECIFICATION RECORDS:  6
001   TIN_S_AG   (1             )  00000       TEACHERS INTERESTED: STRONGLY AGREE
002   TIN__AG    (2             )  10000       TEACHERS INTERESTED: AGREE
003   TIN__UND   (3             )  01000       TEACHERS INTERESTED: UNDECIDED
004   TIN__DAG   (4             )  00100       TEACHERS INTERESTED: DISAGREE
005   TIN_SDAG   (5             )  00010       TEACHERS INTERESTED: STRONGLY DISAGREE
```

BEST COPY AVAILABLE

Table F-4 (continued)
Conditioning Variables for Writing Main Samples

```
006   TIN_MISS   (M          )  00001       TEACHERS INTERESTED: MISSING, OMIT

CONDITIONING ID:             WRIT0001
DESCRIPTION:                 STUDENT: I LIKE TO WRITE
GRADES/ASSESSMENTS:          N04, N08, N12
GROUP LABEL:                 LIKE WRT           LENGTH OF CONTRAST FIELD      :  5
NAEP ID:                     W801601            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS              NUMBER OF SPECIFICATION RECORDS:  6
001   LKW_S_AG   (1          )  00000       STUDENT LIKES TO WRITE: STRONGLY AGREE
002   LKW__AG    (2          )  10000       STUDENT LIKES TO WRITE: AGREE
003   LKW__UND   (3          )  01000       STUDENT LIKES TO WRITE: UNDECIDED
004   LKW__DAG   (4          )  00100       STUDENT LIKES TO WRITE: DISAGREE
005   LKW_SDAG   (5          )  00010       STUDENT LIKES TO WRITE: STRONGLY DISAGREE
006   LKW_MISS   (M          )  00001       STUDENT LIKES TO WRITE: MISSING, OMIT

CONDITIONING ID:             WRIT0002
DESCRIPTION:                 STUDENT SAYS: I AM A GOOD WRITER
GRADES/ASSESSMENTS:          N04, N08, N12
GROUP LABEL:                 GOOD WRT           LENGTH OF CONTRAST FIELD      :  5
NAEP ID:                     W801602            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS              NUMBER OF SPECIFICATION RECORDS:  6
001   GDW_S_AG   (1          )  00000       STUDENT SAYS AM GOOD WRITER: STRONGLY AGREE
002   GDW__AG    (2          )  10000       STUDENT SAYS AM GOOD WRITER: AGREE
003   GDW__UND   (3          )  01000       STUDENT SAYS AM GOOD WRITER: UNDECIDED
004   GDW__DAG   (4          )  00100       STUDENT SAYS AM GOOD WRITER: DISAGREE
005   GDW_SDAG   (5          )  00010       STUDENT SAYS AM GOOD WRITER: STRONGLY DISAGREE
006   GDW_MISS   (M          )  00001       STUDENT SAYS AM GOOD WRITER: MISSING, OMIT

CONDITIONING ID:             WRIT0003
DESCRIPTION:                 HOW OFTEN: WRITE STORY OR REPORT?
GRADES/ASSESSMENTS:          N04
GROUP LABEL:                 WRT STOR           LENGTH OF CONTRAST FIELD      :  4
NAEP ID:                     W801001            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS              NUMBER OF SPECIFICATION RECORDS:  5
001   WST__DAY   (1          )  0000        HOW OFTEN STUDENT WRITES STORY: ALMOST EVERY DAY
002   WST___WK   (2          )  1000        HOW OFTEN STUDENT WRITES STORY: ONCE OR TWICE A WEEK
003   WST__MON   (3          )  0100        HOW OFTEN STUDENT WRITES STORY: ONCE OR TWICE A MONTH
004   WST__NEV   (4          )  0010        HOW OFTEN STUDENT WRITES STORY: NEVER, OR HARDLY EVER
005   WST_MISS   (M          )  0001        HOW OFTEN STUDENT WRITES STORY: MISSING

CONDITIONING ID:             WRIT0004
DESCRIPTION:                 HOW OFTEN: 1-2 PARAGRAPHS ASSIGNED
GRADES/ASSESSMENTS:          N08, N12
GROUP LABEL:                 1-2 PARA           LENGTH OF CONTRAST FIELD      :  4
NAEP ID:                     W801701            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS              NUMBER OF SPECIFICATION RECORDS:  5
001   PAR__DAY   (1          )  0000        HOW OFTEN STUDENT WRITES 1-2 PARAGRAPHS: ALMOST EVERY DAY
002   PAR___WK   (2          )  1000        HOW OFTEN STUDENT WRITES 1-2 PARAGRAPHS: ONCE,TWICE A WEEK
003   PAR__MON   (3          )  0100        HOW OFTEN STUDENT WRITES 1-2 PARAGRAPHS: ONCE,TWICE A MONTH
004   PAR__NEV   (4          )  0010        HOW OFTEN STUDENT WRITES 1-2 PAR: STORY: NEVER/ HARDLY EVER
005   PAR_MISS   (M          )  0001        HOW OFTEN STUDENT WRITES 1-2 PARAGARPHS: STORY: MISSING

CONDITIONING ID:             WRIT0005
DESCRIPTION:                 HOW OFTEN: 1-2 PAGES ASSIGNED
GRADES/ASSESSMENTS:          N08, N12
GROUP LABEL:                 1-2 PAGE           LENGTH OF CONTRAST FIELD      :  4
NAEP ID:                     W801702            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS              NUMBER OF SPECIFICATION RECORDS:  5
001   PGS__DAY   (1          )  0000        HOW OFTEN STUDENT WRITES 1-2 PAGES: ALMOST EVERY DAY
002   PGS___WK   (2          )  1000        HOW OFTEN STUDENT WRITES 1-2 PAGES: ONCE OR TWICE A WEEK
003   PGS__MON   (3          )  0100        HOW OFTEN STUDENT WRITES 1-2 PAGES: ONCE OR TWICE A MONTH
004   PGS__NEV   (4          )  0010        HOW OFTEN STUDENT WRITES 1-2 PAGES: NEVER, OR HARDLY EVER
005   PGS_MISS   (M          )  0001        HOW OFTEN STUDENT WRITES 1-2 PAGES: MISSING

CONDITIONING ID:             WRIT0006
```

Table F-4 (continued)
Conditioning Variables for Writing Main Samples

```
DESCRIPTION:                HOW OFTEN: 3 OR MORE PAGES ASSIGNED
GRADES/ASSESSMENTS:         N08, N12
GROUP LABEL:                3+ PAGES            LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                    W801703             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   5
001   3PP__DAY   (1            )  0000          HOW OFTEN STUDENT WRITES 3+ PAGES: ALMOST EVERY DAY
002   3PP___WK   (2            )  1000          HOW OFTEN STUDENT WRITES 3+ PAGES: ONCE OR TWICE A WEEK
003   3PP__MON   (3            )  0100          HOW OFTEN STUDENT WRITES 3+ PAGES: ONCE OR TWICE A MONTH
004   3PP__NEV   (4            )  0010          HOW OFTEN STUDENT WRITES 3+ PAGES: NEVER, OR HARDLY EVER
005   3PP_MISS   (M            )  0001          HOW OFTEN STUDENT WRITES 3+ PAGES: MISSING

CONDITIONING ID:            WRIT0007
DESCRIPTION:                HOW OFTEN: REPORT OR SUMMARY
GRADES/ASSESSMENTS:         N08, N12
GROUP LABEL:                REPORTS             LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                    W801801             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   5
001   REP__DAY   (1            )  0000          HOW OFTEN STUDENT WRITES REPORT: ALMOST EVERY DAY
002   REP___WK   (2            )  1000          HOW OFTEN STUDENT WRITES REPORT: ONCE OR TWICE A WEEK
003   REP__MON   (3            )  0100          HOW OFTEN STUDENT WRITES REPORT: ONCE OR TWICE A MONTH
004   REP__NEV   (4            )  0010          HOW OFTEN STUDENT WRITES REPORT: NEVER, OR HARDLY EVER
005   REP_MISS   (M            )  0001          HOW OFTEN STUDENT WRITES REPORT: MISSING

CONDITIONING ID:            WRIT0008
DESCRIPTION:                HOW OFTEN: ANALYSIS OR INTERPRETATION
GRADES/ASSESSMENTS:         N08, N12
GROUP LABEL:                ANALYSES            LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                    W801802             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   5
001   ANL__DAY   (1            )  0000          HOW OFTEN STUDENT WRITES ANALYSES: ALMOST EVERY DAY
002   ANL___WK   (2            )  1000          HOW OFTEN STUDENT WRITES ANALYSES: ONCE OR TWICE A WEEK
003   ANL__MON   (3            )  0100          HOW OFTEN STUDENT WRITES ANALYSES: ONCE OR TWICE A MONTH
004   ANL__NEV   (4            )  0010          HOW OFTEN STUDENT WRITES ANALYSES: NEVER, OR HARDLY EVER
005   ANL_MISS   (M            )  0001          HOW OFTEN STUDENT WRITES ANALYSES: MISSING

CONDITIONING ID:            WRIT0009
DESCRIPTION:                HOW OFTEN: WRITE  PERSUASIVE ESSAY
GRADES/ASSESSMENTS:         N08, N12
GROUP LABEL:                PERS ESS            LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                    W801803             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   5
001   PER__DAY   (1            )  0000          HOW OFTEN STUDENT WRITES PERSUASIVE ESSAY: ALMOST EVERY DAY
002   PER___WK   (2            )  1000          HOW OFTEN STUDENT WRITES PERSUASIVE ESSAY: ONCE,TWICE WEEK
003   PER__MON   (3            )  0100          HOW OFTEN STUDENT WRITES PERSUASIVE ESSAY: ONCE,TWICE MONTH
004   PER__NEV   (4            )  0010          HOW OFTEN STUDENT WRITES PERSUASIVE ESSAY: NEVER,HARDLY EVER
005   PER_MISS   (M            )  0001          HOW OFTEN STUDENT WRITES PERSUASIVE ESSAY: MISSING

CONDITIONING ID:            WRIT0010
DESCRIPTION:                HOW OFTEN: WRITE PERSONAL/IMAGINED NARRATIVE
GRADES/ASSESSMENTS:         N08, N12
GROUP LABEL:                NARRATIV            LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                    W801804             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   5
001   NAR__DAY   (1            )  0000          HOW OFTEN STUDENT WRITES NARRATIVE: ALMOST EVERY DAY
002   NAR___WK   (2            )  1000          HOW OFTEN STUDENT WRITES NARRATIVE: ONCE OR TWICE A WEEK
003   NAR__MON   (3            )  0100          HOW OFTEN STUDENT WRITES NARRATIVE: ONCE OR TWICE A MONTH
004   NAR__NEV   (4            )  0010          HOW OFTEN STUDENT WRITES NARRATIVE: NEVER, OR HARDLY EVER
005   NAR_MISS   (M            )  0001          HOW OFTEN STUDENT WRITES NARRATIVE: MISSING

CONDITIONING ID:            WRIT0011
DESCRIPTION:                HOW OFTEN: PLAN YOUR WRITING
GRADES/ASSESSMENTS:         N08, N12
GROUP LABEL:                PLAN_WRT            LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                    W801204             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   4
```

Table F-4 (continued)
Conditioning Variables for Writing Main Samples

```
001   PLA__ALW   (1          )  000        HOW OFTEN STUDENT PLAN WRITING: ALWAYS
002   PLA_SOME   (2          )  100        HOW OFTEN STUDENT PLAN WRITING: SOMETIMES
003   PLA__NEV   (3          )  010        HOW OFTEN STUDENT PLAN WRITING: NEVER
004   PLA_MISS   (M          )  001        HOW OFTEN STUDENT PLAN WRITING: MISSING

CONDITIONING ID:            WRIT0012
DESCRIPTION:                HOW OFTEN: MAKE AN OUTLINE BEFORE WRITING
GRADES/ASSESSMENTS:         N08, N12
GROUP LABEL:                OUTLINE             LENGTH OF CONTRAST FIELD      :  3
NAEP ID:                    W801205             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   OTL__ALW   (1           )  000       HOW OFTEN STUDENT OUTLINE BEFORE WRITING: ALWAYS
002   OTL_SOME   (2           )  100       HOW OFTEN STUDENT OUTLINE BEFORE WRITING: SOMETIMES
003   OTL__NEV   (3           )  010       HOW OFTEN STUDENT OUTLINE BEFORE WRITING: NEVER
004   OTL_MISS   (M           )  001       HOW OFTEN STUDENT OUTLINE BEFORE WRITING: MISSING

CONDITIONING ID:            WRIT0013
DESCRIPTION:                HOW OFTEN: DEFINE PURPOSE OR AUDIENCE
GRADES/ASSESSMENTS:         N08, N12
GROUP LABEL:                DEF_AUD             LENGTH OF CONTRAST FIELD      :  3
NAEP ID:                    W801206             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   AUD__ALW   (1           )  000       HOW OFTEN STUDENT DEFINE PURPOSE OR AUDIENCE: ALWAYS
002   AUD_SOME   (2           )  100       HOW OFTEN STUDENT DEFINE PURPOSE OR AUDIENCE: SOMETIMES
003   AUD__NEV   (3           )  010       HOW OFTEN STUDENT DEFINE PURPOSE OR AUDIENCE: NEVER
004   AUD_MISS   (M           )  001       HOW OFTEN STUDENT DEFINE PURPOSE OR AUDIENCE: MISSING

CONDITIONING ID:            WRIT0014
DESCRIPTION:                HOW OFTEN: TALK ABOUT PAPER WHILE WRITING
GRADES/ASSESSMENTS:         N04, N08, N12
GROUP LABEL:                TALK_WRT            LENGTH OF CONTRAST FIELD      :  3
NAEP ID:                    W801201             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   TLK__ALW   (1           )  000       HOW OFTEN TALK ABOUT PAPER WHILE WRITING: ALWAYS
002   TLK_SOME   (2           )  100       HOW OFTEN TALK ABOUT PAPER WHILE WRITING: SOMETIMES
003   TLK__NEV   (3           )  010       HOW OFTEN TALK ABOUT PAPER WHILE WRITING: NEVER
004   TLK_MISS   (M           )  001       HOW OFTEN TALK ABOUT PAPER WHILE WRITING: MISSING

CONDITIONING ID:            WRIT0015
DESCRIPTION:                HOW OFTEN: STUDENT USES SOURCES OTHER THAN TEXT
GRADES/ASSESSMENTS:         N08, N12
GROUP LABEL:                SOURCES             LENGTH OF CONTRAST FIELD      :  3
NAEP ID:                    W801207             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   SOU__ALW   (1           )  000       HOW OFTEN STUDENT USES SOURCES OTHER THAN TEXT: ALWAYS
002   SOU_SOME   (2           )  100       HOW OFTEN STUDENT USES SOURCES OTHER THAN TEXT: SOMETIMES
003   SOU__NEV   (3           )  010       HOW OFTEN STUDENT USES SOURCES OTHER THAN TEXT: NEVER
004   SOU_MISS   (M           )  001       HOW OFTEN STUDENT USES SOURCES OTHER THAN TEXT: MISSING

CONDITIONING ID:            WRIT0016
DESCRIPTION:                HOW OFTEN: STUDENT WRITE MORE THAN ONE DRAFT OF PAPER
GRADES/ASSESSMENTS:         N04, N08, N12
GROUP LABEL:                DRAFTS              LENGTH OF CONTRAST FIELD      :  3
NAEP ID:                    W801202             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   DFT__ALW   (1           )  000       HOW OFTEN STUDENT WRITES MORE THAN ONE DRAFT: ALWAYS
002   DFT_SOME   (2           )  100       HOW OFTEN STUDENT WRITES MORE THAN ONE DRAFT: SOMETIMES
003   DFT__NEV   (3           )  010       HOW OFTEN STUDENT WRITES MORE THAN ONE DRAFT: NEVER
004   DFT_MISS   (M           )  001       HOW OFTEN STUDENT WRITES MORE THAN ONE DRAFT: MISSING

CONDITIONING ID:            WRIT0017
DESCRIPTION:                HOW OFTEN: STUDENT CONTRIBUTE WRITING TO A COLLECTION
GRADES/ASSESSMENTS:         N04, N08, N12
GROUP LABEL:                COLLECT             LENGTH OF CONTRAST FIELD      :  3
NAEP ID:                    W801203             DEGREES OF FREEDOM PER CONTRAST:  1
```

Table F-4 (continued)
Conditioning Variables for Writing Main Samples

```
TYPE OF CONTRAST:          CLASS              NUMBER OF SPECIFICATION RECORDS:   4
001   CLT__ALW   (1           )  000          HOW OFTEN STUDENT CONTRIBUTES TO A COLLECTION: ALWAYS
002   CLT_SOME   (2           )  100          HOW OFTEN STUDENT CONTRIBUTES TO A COLLECTION: SOMETIMES
003   CLT__NEV   (3           )  010          HOW OFTEN STUDENT CONTRIBUTES TO A COLLECTION: NEVER
004   CLT_MISS   (M           )  001          HOW OFTEN STUDENT CONTRIBURES TO A COLLECTION: MISSING

CONDITIONING ID:           WRIT0018
DESCRIPTION:               HOW OFTEN: DO SPELLING/PUNCUTATION/GRAMMAR EXERCISES
GRADES/ASSESSMENTS:        N04, N08, N12
GROUP LABEL:               SP/GRAM            LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                   W801101            DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS              NUMBER OF SPECIFICATION RECORDS:   5
001   SPG__DAY   (1           )  0000         HOW OFTEN DOES SPELL/PUNC/GRAM EXERCISES: ALMOST EVERY DAY
002   SPG___WK   (2           )  1000         HOW OFTEN DOES SPELL/PUNC/GRAM EXERCISES: ONCE OR TWICE/WK
003   SPG__MON   (3           )  0100         HOW OFTEN DOES SPELL/PUNC/GRAM EXERCISES: ONCE OR TWICE/MON
004   SPG__NEV   (4           )  0010         HOW OFTEN DOES SPELL/PUNC/GRAM EXERCISES: NEVER/HARDLY EVER
005   SPG_MISS   (M           )  0001         HOW OFTEN DOES SPELLING/PUNC/GRAMMAR EXERCISES: MISSING

CONDITIONING ID:           WRIT0019
DESCRIPTION:               HOW OFTEN: STUDENT WORK IN GROUPS TO DISCUSS WRITING
GRADES/ASSESSMENTS:        N04, N08, N12
GROUP LABEL:               GROUPS             LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                   W801102            DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS              NUMBER OF SPECIFICATION RECORDS:   5
001   GRP__DAY   (1           )  0000         HOW OFTEN WORK IN GROUPS TO DISCUSS WRTNG: ALMOST EVERY DAY
002   GRP___WK   (2           )  1000         HOW OFTEN WORK IN GROUPS TO DISCUSS WRITING: ONCE/TWICE/WK
003   GRP__MON   (3           )  0100         HOW OFTEN WORK IN GROUPS TO DISCUSS WRITING: ONCE/TWICE/MON
004   GRP__NEV   (4           )  0010         HOW OFTEN WORK IN GROUPS TO DISCUSS WRITING: NEVER,HARDLY
005   GRP_MISS   (M           )  0001         HOW OFTEN STUDENT WORK IN GROUPS TO DISCUSS WRITING: MISSING

CONDITIONING ID:           WRIT0020
DESCRIPTION:               HOW OFTEN: STUDENT WRITE IN A LOG/JOURNAL
GRADES/ASSESSMENTS:        N04, N08, N12
GROUP LABEL:               LOG/JOUR           LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                   W801103            DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS              NUMBER OF SPECIFICATION RECORDS:   5
001   LOG__DAY   (1           )  0000         HOW OFTEN STUDENT WRITE IN LOG/JOURNAL: ALMOST EVERY DAY
002   LOG___WK   (2           )  1000         HOW OFTEN STUDENT WRITE IN LOG/JOURNAL: ONCE OR TWICE A WEEK
003   LOG__MON   (3           )  0100         HOW OFTEN STUDENT WRITE IN LOG/JOURNAL: ONCE OR TWICE/MONTH
004   LOG__NEV   (4           )  0010         HOW OFTEN STUDENT WRITE IN LOG/JOURNAL: NEVER, HARDLY EVER
005   LOG_MISS   (M           )  0001         HOW OFTEN STUDENT WRITE IN LOG/JOURNAL: MISSING

CONDITIONING ID:           WRIT0021
DESCRIPTION:               IS WRITING KEPT IN A PORTFOLIO?
GRADES/ASSESSMENTS:        N04, N08, N12
GROUP LABEL:               PORTFOL            LENGTH OF CONTRAST FIELD       :   2
NAEP ID:                   W801301            DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS              NUMBER OF SPECIFICATION RECORDS:   3
001   PORTF__Y   (1           )  00           IS WRITING KEPT IN A PORTFOLIO: YES
002   PORTF__N   (2           )  10           IS WRITING KEPT IN A PORTFOLIO: NO
003   PORTF__?   (M           )  01           IS WRITING KEPT IN A PORTFOLIO: MISSING

CONDITIONING ID:           WRIT0022
DESCRIPTION:               WHEN TEACHER GRADES, HOW IMPORTANT IS SPELLING/GRAMMAR
GRADES/ASSESSMENTS:        N04, N08, N12
GROUP LABEL:               IMPT SPG           LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                   W801401            DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS              NUMBER OF SPECIFICATION RECORDS:   4
001   ISG__ALW   (1           )  000          HOW IMPORTANT IS SPELLING/GRAMMAR: VERY IMPORTANT
002   ISG_SOME   (2           )  100          HOW IMPORTANT IS SPELLING/GRAMMAR: MODERATELY IMPORTAN
003   ISG__NEV   (3           )  010          HOW IMPORTANT IS SPELLING/GRAMMAR: NOT VERY IMPORTANT
004   ISG_MISS   (M           )  001          HOW IMPORTANT IS SPELLING/GRAMMAR: MISSING

CONDITIONING ID:           WRIT0023
DESCRIPTION:               WHEN TEACHER GRADES, HOW IMPORTANT IS ORGANIZED PAPER
```

Table F-4 (continued)
Conditioning Variables for Writing Main Samples

```
GRADES/ASSESSMENTS:       N04, N08, N12
GROUP LABEL:              IMPT ORG           LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                  W801402            DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:         CLASS              NUMBER OF SPECIFICATION RECORDS:   4
001   IORG_ALW   (1          )  000          HOW IMPORTANT IS ORGANIZED PAPER: VERY IMPORTANT
002   IORG_SOM   (2          )  100          HOW IMPORTANT IS ORGANIZED PAPER: MODERATELY IMPORTANT
003   IORG_NEV   (3          )  010          HOW IMPORTANT IS ORGANIZED PAPER: NOT VERY IMPORTANT
004   IORG_MIS   (M          )  001          HOW IMPORTANT IS ORGANIZED PAPER: MISSING

CONDITIONING ID:          WRIT0024
DESCRIPTION:              WHEN TEACHER GRADES, HOW IMPORTANT IS QUALITY/CREATIVITY
GRADES/ASSESSMENTS:       N04, N08, N12
GROUP LABEL:              IMPT CRE           LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                  W801403            DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:         CLASS              NUMBER OF SPECIFICATION RECORDS:   4
001   ICRE_ALW   (1          )  000          HOW IMPORTANT IS QUALITY/CREATIVITY: VERY IMPORTANT
002   ICRE_SOM   (2          )  100          HOW IMPORTANT IS QUALITY/CREATIVITY: MODERATELY IMPORT
003   ICRE_NEV   (3          )  010          HOW IMPORTANT IS QUALITY/CREATIVITY: NOT VERY IMPORTAN
004   ICRE_MIS   (M          )  001          HOW IMPORTANT IS QUALITY/CREATICITY: MISSING

CONDITIONING ID:          WRIT0025
DESCRIPTION:              WHEN TEACHER GRADES, HOW IMPORTANT IS LENGTH
GRADES/ASSESSMENTS:       N04, N08, N12
GROUP LABEL:              IMPT LEN           LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                  W801404            DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:         CLASS              NUMBER OF SPECIFICATION RECORDS:   4
001   ILEN_ALW   (1          )  000          HOW IMPORTANT IS LENGTH: VERY IMPORTANT
002   ILEN_SOM   (2          )  100          HOW IMPORTANT IS LENGTH: MODERATELY IMPORTANT
003   ILEN_NEV   (3          )  010          HOW IMPORTANT IS LENGTH: NOT VERY IMPORTANT
004   ILEN_MIS   (M          )  001          HOW IMPORTANT IS LENGTH: MISSING

CONDITIONING ID:          WRIT0026
DESCRIPTION:              HOW OFTEN ON A COMPUTER DO YOU DO SPELLING/PUNCUTATION/GRAMMAR EXERCISES
GRADES/ASSESSMENTS:       N04, N08, N12
GROUP LABEL:              COMP SG            LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                  W801501            DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:         CLASS              NUMBER OF SPECIFICATION RECORDS:   4
001   CSP___WK   (1          )  000          HOW OFTEN ON COMPUTER DOES SPELL/PUNC/GRAMMAR: ONCE/WK
002   CSP__MON   (2          )  100          HOW OFTEN ON COMPUTER DOES SPELL/PUNC/GRAMMAR: ONCE/MON
003   CSP__NEV   (3          )  010          HOW OFTEN ON COMPUTER DOES SPELL/PUNC/GRAMMAR: NEVER
004   CSP_MISS   (M          )  001          HOW OFTEN ON COMPUTER DOES SPELL/PUNC/GRAMMAR: MISSING

CONDITIONING ID:          WRIT0027
DESCRIPTION:              HOW OFTEN ON A COMPUTER WRITE STORIES OR REPORTS
GRADES/ASSESSMENTS:       N04, N08, N12
GROUP LABEL:              COMP REP           LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                  W801502            DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:         CLASS              NUMBER OF SPECIFICATION RECORDS:   4
001   CRP___WK   (1          )  000          HOW OFTEN STUDENT WRITES ON COMPUTER: ONCE OR TWICE WEEK
002   CRP__MON   (2          )  100          HOW OFTEN STUDENT WRITES ON COMPUTER: ONCE OR TWICE MONTH
003   CRP__NEV   (3          )  010          HOW OFTEN STUDENT WRITES ON COMPUTER: NEVER, HARDLY EVER
004   CRP_MISS   (M          )  001          HOW OFTEN STUDENT WRITES ON COMPUTER: MISSING

CONDITIONING ID:          WRIT0028
DESCRIPTION:              HOW WELL DO YOU THINK YOU DID ON THIS (WRITING) TEST
GRADES/ASSESSMENTS:       N04, S04, N08, S08, N12
GROUP LABEL:              U_DO_WEL           LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                  WM00101            DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:         CLASS              NUMBER OF SPECIFICATION RECORDS:   4
001   DOWEL_VW   (1          )  000          HOW WELL DID YOU DO:  VERY WELL
002   DOWEL_PG   (2          )  100          HOW WELL DID YOU DO:  PRETTY WELL
003   DOWEL_NG   (3          )  010          HOW WELL DID YOU DO:  NOT VERY WELL
004   DOWEL_BD   (4,M        )  001          HOW WELL DID YOU DO:  NOT WELL AT ALL, MISSING

CONDITIONING ID:          WRIT0029
```

Table F-4 (continued)
Conditioning Variables for Writing Main Samples

```
DESCRIPTION:                  HOW HARD WAS THIS TEST COMPARED TO OTHERS?
GRADES/ASSESSMENTS:           N04, S04, N08, S08, N12
GROUP LABEL:                  TEST_DIF          LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                      WM00201           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:             CLASS             NUMBER OF SPECIFICATION RECORDS:   4
001   TESTDIF1   (1,M           )  000          TEST DIFFICULTY:  MUCH HARDER THAN OTHERS
002   TESTDIF2   (2             )  100          TEST DIFFICULTY:  HARDER THAN OTHERS
003   TESTDIF3   (3             )  010          TEST DIFFICULTY:  ABOUT AS HARD AS OTHERS
004   TESTDIF4   (4             )  001          TEST DIFFICULTY:  EASIER THAN OTHERS

CONDITIONING ID:              WRIT0030
DESCRIPTION:                  HOW HARD DID YOU TRY ON THIS TEST COMPRAED TO OTHERS?
GRADES/ASSESSMENTS:           N04, S04, N08, S08, N12
GROUP LABEL:                  TEST_EFF          LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                      WM00301           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:             CLASS             NUMBER OF SPECIFICATION RECORDS:   4
001   TESTEFF1   (1,M           )  000          TEST EFFORT:  MUCH HARDER THAN OTHERS
002   TESTEFF2   (2             )  100          TEST EFFORT:  HARDER THAN OTHERS
003   TESTEFF3   (3             )  010          TEST EFFORT:  ABOUT AS HARD AS OTHERS
004   TESTEFF4   (4             )  001          TEST EFFORT:  NOT AS HARD AS OTHERS

CONDITIONING ID:              WRIT0031
DESCRIPTION:                  HOW IMPORTANT WAS IT TO YOU TO DO WELL?
GRADES/ASSESSMENTS:           N04, S04, N08, S08, N12
GROUP LABEL:                  TEST_IMP          LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                      WM00401           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:             CLASS             NUMBER OF SPECIFICATION RECORDS:   5
001   TESTIMP1   (1             )  0000         TEST IMPORTANCE:  VERY IMPORTANT
002   TESTIMP2   (2             )  1000         TEST IMPORTANCE:  IMPORTANT
003   TESTIMP3   (3             )  0100         TEST IMPORTANCE:  SOMEWHAT IMPORTANT
004   TESTIMP4   (4             )  0010         TEST IMPORTANCE:  NOT VERY IMPORTANT
005   TESTIMP?   (M             )  0001         TEST IMPORTANCE:  MISSING

CONDITIONING ID:              WRIT0032
DESCRIPTION:                  HOW OFTEN WERE YOU ASKED TO WRITE ESSAYS FOR WRITING PORTFOLIO?
GRADES/ASSESSMENTS:           N04, S04, N08, S08, N12
GROUP LABEL:                  ESSPORT           LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                      WM00501           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:             CLASS             NUMBER OF SPECIFICATION RECORDS:   4
001   ESSPOR_W   (1             )  000          DETAILED SOLUTIONS:  AT LEAST ONCE A WEEK
002   ESSPOR_M   (2             )  100          DETAILED SOLUTIONS:  ONCE OR TWICE A MONTH
003   ESSPOR_Y   (3             )  010          DETAILED SOLUTIONS:  ONCE OR TWICE A YEAR
004   ESSPOR_N   (4,M           )  001          DETAILED SOLUTIONS:  NEVER, MISSING

CONDITIONING ID:              SCHL0001
DESCRIPTION:                  WHO TEACHES ENGLISH/LANGUAGE ARTS TO 8TH GRADE?
GRADES/ASSESSMENTS:           N08, S08
GROUP LABEL:                  TSUB_ENG          LENGTH OF CONTRAST FIELD       :   2
NAEP ID:                      C034701           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:             CLASS             NUMBER OF SPECIFICATION RECORDS:   3
001   ENG-T>1S   (1,M           )  00           8TH-GRADE ENGLISH:  TEACHERS WITH MORE THAN ONE SUBJECT
002   ENG-T=1S   (2             )  10           8TH-GRADE ENGLISH:  TEACHERS WITH ONE SUBJECT
003   ENG-SNT    (3             )  01           8TH-GRADE ENGLISH:  SUBJECT NOT TAUGHT

CONDITIONING ID:              SCHL0002
DESCRIPTION:                  HAS READING BEEN IDENTIFIED AS A PRIORITY? (GRADE 4)
GRADES/ASSESSMENTS:           N04, S04, N08, S08
GROUP LABEL:                  PRIOR-RD          LENGTH OF CONTRAST FIELD       :   2
NAEP ID:                      C031601           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:             CLASS             NUMBER OF SPECIFICATION RECORDS:   3
001   RPRIOR-Y   (1             )  00           READING PRIORI
002   RPRIOR-N   (2             )  10           READING PRIO
003   RPRIOR-?   (M             )  01           READING PRIORITY:  MISSING

CONDITIONING ID:              SCHL0003
```

Table F-4 (continued)
Conditioning Variables for Writing Main Samples

```
DESCRIPTION:                HAS WRITING BEEN IDENTIFIED AS A PRIORITY? (GRADE 4)
GRADES/ASSESSMENTS:         N04, S04, N08, S08
GROUP LABEL:                PRIOR-WR          LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                    C031602           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS             NUMBER OF SPECIFICATION RECORDS:  3
001   WPRIOR-Y    (1          )  00           WRITING PRIORITY:  YES
002   WPRIOR-N    (2          )  10           WRITING PRIORITY:  NO
003   WPRIOR-?    (M          )  01           WRITING PRIORITY:  MISSING

CONDITIONING ID:            SCHL0004
DESCRIPTION:                WHAT PERCENT OF STUDENTS RECEIVE SUBSIDIZED LUNCH?
GRADES/ASSESSMENTS:         N04, S04, N08, S08, N12
GROUP LABEL:                %SUBLUN           LENGTH OF CONTRAST FIELD       :  5
NAEP ID:                    C032001           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS             NUMBER OF SPECIFICATION RECORDS:  6
001   %SUBLUN1    (1,2,3      )  00000        PERCENT SUBSIDIZED LUNCH:  NONE-10%
002   %SUBLUN2    (4          )  10000        PERCENT SUBSIDIZED LUNCH:  11-25%
003   %SUBLUN3    (5          )  01000        PERCENT SUBSIDIZED LUNCH:  26-50%
004   %SUBLUN4    (6          )  00100        PERCENT SUBSIDIZED LUNCH:  51-75%
005   %SUBLUN5    (7,8        )  00010        PERCENT SUBSIDIZED LUNCH:  76-100%
006   %SUBLUN?    (M          )  00001        PERCENT SUBSIDIZED LUNCH:  MISSING

CONDITIONING ID:            SCHL0005
DESCRIPTION:                WHAT PERCENT OF STUDENTS RECEIVE REMEDIAL READING?
GRADES/ASSESSMENTS:         N04, S04, N08, S08, N12
GROUP LABEL:                %REMDL-R          LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                    C032002           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS             NUMBER OF SPECIFICATION RECORDS:  5
001   %REMRED1    (1,2        )  0000         PERCENT REMEDIAL READING:  NONE-5%
002   %REMRED2    (3          )  1000         PERCENT REMEDIAL READING:  6-10%
003   %REMRED3    (4          )  0100         PERCENT REMEDIAL READING:  11-25%
004   %REMRED4    (5,6,7,8    )  0010         PERCENT REMEDIAL READING:  26-100%
005   %REMRED?    (M          )  0001         PERCENT REMEDIAL READING:  MISSING

CONDITIONING ID:            SCHL0010
DESCRIPTION:                ARE 8TH GRADERS ASSIGNED TO ENGLISH BY ABILITY?
GRADES/ASSESSMENTS:         N08, S08
GROUP LABEL:                ENG/AB8           LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                    C034401           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS             NUMBER OF SPECIFICATION RECORDS:  3
001   ENG/AB-Y    (1          )  00           ENGLISH BY ABILITY:  YES
002   ENG/AB-N    (2          )  10           ENGLISH BY ABILITY:  NO
003   ENG/AB-?    (M          )  01           ENGLISH BY ABILITY:  MISSING

CONDITIONING ID:            SCHL0011
DESCRIPTION:                ARE 12TH GRADERS ASSIGNED TO ENGLISH BY ABILITY?
GRADES/ASSESSMENTS:         N12
GROUP LABEL:                ENG/AB12          LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                    C035001           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS             NUMBER OF SPECIFICATION RECORDS:  3
001   ENG/AB-Y    (1          )  00           ENGLISH BY ABILITY:  YES
002   ENG/AB-N    (2          )  10           ENGLISH BY ABILITY:  NO
003   ENG/AB-?    (M          )  01           ENGLISH BY ABILITY:  MISSING

CONDITIONING ID:            SCHL0012
DESCRIPTION:                NUMBER OF SEMESTERS ENGLISH/LITERATURE/WRITING REQUIRED (GRADE 12)
GRADES/ASSESSMENTS:         N12
GROUP LABEL:                S_SEMENG          LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                    C035201           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS             NUMBER OF SPECIFICATION RECORDS:  3
001   #SEMENG0    (1-8        )  00           SEMESTERS ENGLISH:  NONE-SEVEN
002   #SEMENG8    (9          )  10           SEMESTERS ENGLISH:  EIGHT
003   #SEMENG?    (M          )  01           SEMESTERS ENGLISH:  MISSING

CONDITIONING ID:            SCHL0014
```

Table F-4 (continued)
Conditioning Variables for Writing Main Samples

```
DESCRIPTION:              POLICY CONTROLLING TIME FOR READING INSTRUCTION?
GRADES/ASSESSMENTS:       N04, S04
GROUP LABEL:              POLICY-R           LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                  C031301            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS              NUMBER OF SPECIFICATION RECORDS:  3
001   RD_POL-Y   (1          )  00           READING TIME POLICY:  YES
002   RD_POL-N   (2          )  10           READING TIME POLICY:  NO
003   RD_POL-?   (M          )  01           READING TIME POLICY:  MISSING

CONDITIONING ID:          SCHL0015
DESCRIPTION:              POLICY CONTROLLING TIME FOR WRITING INSTRUCTION?
GRADES/ASSESSMENTS:       N04, S04
GROUP LABEL:              POLICY-W           LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                  C031302            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS              NUMBER OF SPECIFICATION RECORDS:  3
001   WR_POL-Y   (1          )  00           WRITING TIME POLICY:  YES
002   WR_POL-N   (2          )  10           WRITING TIME POLICY:  NO
003   WR_POL-?   (M          )  01           WRITING TIME POLICY:  MISSING

CONDITIONING ID:          SCHL0016
DESCRIPTION:              ARE COMPUTERS ALWAYS AVAILABLE IN ENGLISH CLASSROOMS? (GRADE 8)
GRADES/ASSESSMENTS:       N08, S08, N12
GROUP LABEL:              COMP_ECL           LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                  C035601            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS              NUMBER OF SPECIFICATION RECORDS:  3
001   CMPECL-Y   (1          )  00           COMPUTERS AVAILABLE IN ENGLISH CLASS:  YES
002   CMPECL-N   (2          )  10           COMPUTERS AVAILABLE IN ENGLISH CLASS:  NO
003   CMPECL-?   (M          )  01           COMPUTERS AVAILABLE IN ENGLISH CLASS:  MISSING

CONDITIONING ID:          SCHL0017
DESCRIPTION:              ARE COMPUTERS ALWAYS GROUPED IN A LAB AVAILABLE FOR ENGLISH CLASSES? (GRADE 8)
GRADES/ASSESSMENTS:       N08, S08, N12
GROUP LABEL:              COMP_ELB           LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                  C035602            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS              NUMBER OF SPECIFICATION RECORDS:  3
001   CMPELB-Y   (1          )  00           COMPUTERS IN LAB FOR ENGLISH CLASS:  YES
002   CMPELB-N   (2          )  10           COMPUTERS IN LAB FOR ENGLISH CLASS:  NO
003   CMPELB-?   (M          )  01           COMPUTERS IN LAB FOR ENGLISH CLASS:  MISSING

CONDITIONING ID:          SCHL0018
DESCRIPTION:              ARE COMPUTERS AVAILABLE TO BRING TO ENGLISH CLASSES? (GRADE 8)
GRADES/ASSESSMENTS:       N08, S08, N12
GROUP LABEL:              COMP_EBR           LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                  C035603            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS              NUMBER OF SPECIFICATION RECORDS:  3
001   CMPEBR-Y   (1          )  00           COMPUTERS AVAILABLE IN ENGLISH CLASS:  YES
002   CMPEBR-N   (2          )  10           COMPUTERS AVAILABLE IN ENGLISH CLASS:  NO
003   CMPEBR-?   (M          )  01           COMPUTERS AVAILABLE IN ENGLISH CLASS:  MISSING

CONDITIONING ID:          SCHL0019
DESCRIPTION:              NUMBER OF STUDENTS ENROLLED IN AP ENGLISH
GRADES/ASSESSMENTS:       N12
GROUP LABEL:              #AP_ENGL           LENGTH OF CONTRAST FIELD       :  6
NAEP ID:                  C035802            DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS              NUMBER OF SPECIFICATION RECORDS:  7
001   AP_ENGL1   (1          )  000000       # STUDENTS AP ENGLISH:  NONE
002   AP_ENGL2   (2          )  100000       # STUDENTS AP ENGLISH:  1-5
003   AP_ENGL3   (3          )  010000       # STUDENTS AP ENGLISH:  6-10
004   AP_ENGL4   (4          )  001000       # STUDENTS AP ENGLISH:  11-25
005   AP_ENGL5   (5          )  000100       # STUDENTS AP ENGLISH:  26-50
006   AP_ENGL6   (6          )  000010       # STUDENTS AP ENGLISH:  MORE THAN 50
007   AP_ENGL?   (M          )  000001       # STUDENTS AP ENGLISH:  MISSING

CONDITIONING ID:          SCHL0020
DESCRIPTION:              DOES SCHOOL INVOLVE PARENTS AS AIDES IN CLASS?
```

Table F-4 (continued)
Conditioning Variables for Writing Main Samples

```
GRADES/ASSESSMENTS:       N04, S04, N08, S08, N12
GROUP LABEL:              PAR_AIDE          LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                  C032207           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  4
001   PARAID-R   (1         )  000          PARENTS AS AIDES IN CLASS:  ROUTINELY
002   PARAID-O   (2         )  100          PARENTS AS AIDES IN CLASS:  OCCASIONALLY
003   PARAID-N   (3         )  010          PARENTS AS AIDES IN CLASS:  NO
004   PARAID-?   (M         )  001          PARENTS AS AIDES IN CLASS:  MISSING

CONDITIONING ID:          SCHL0035
DESCRIPTION:              HOW OFTEN IS 8TH GRADER INSTRUCTED IN COMPUTERS?
GRADES/ASSESSMENTS:       N08, S08
GROUP LABEL:              CMP8_INS          LENGTH OF CONTRAST FIELD       :  5
NAEP ID:                  C034504           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  6
001   CMP_INS1   (1         )  00000        8TH GRADE COMPUTER INSTRUCTION:  EVERY DAY
002   CMP_INS2   (2         )  10000        8TH GRADE COMPUTER INSTRUCTION:  3-4 TIMES A WEEK
003   CMP_INS3   (3         )  01000        8TH GRADE COMPUTER INSTRUCTION:  1-2 TIMES A WEEK
004   CMP_INS4   (4         )  00100        8TH GRADE COMPUTER INSTRUCTION:  LESS THAN ONCE A WEEK
005   CMP_INS5   (5         )  00010        8TH GRADE COMPUTER INSTRUCTION:  NOT TAUGHT
006   CMP_INS?   (M         )  00001        8TH GRADE COMPUTER INSTRUCTION:  MISSING

CONDITIONING ID:          SCHL0077
DESCRIPTION:              ARE COMPUTERS ALWAYS AVAILABLE IN CLASSROOMS? (GRADE 4)
GRADES/ASSESSMENTS:       N04, S04
GROUP LABEL:              COMP_CLA          LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                  C035701           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  3
001   CMPCLA-Y   (1         )  00           COMPUTERS AVAILABLE IN CLASS:  YES
002   CMPCLA-N   (2         )  10           COMPUTERS AVAILABLE IN CLASS:  NO
003   CMPCLA-?   (M         )  01           COMPUTERS AVAILABLE IN CLASS:  MISSING

CONDITIONING ID:          SCHL0078
DESCRIPTION:              ARE COMPUTERS GROUPED IN A COMPUTER LAB? (GRADE 4)
GRADES/ASSESSMENTS:       N04, S04
GROUP LABEL:              COMP_LAB          LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                  C035702           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  3
001   CMPLAB-Y   (1         )  00           COMPUTERS IN A LAB:  YES
002   CMPLAB-N   (2         )  10           COMPUTERS IN A LAB:  NO
003   CMPLAB-?   (M         )  01           COMPUTERS IN A LAB:  MISSING

CONDITIONING ID:          SCHL0079
DESCRIPTION:              ARE COMPUTERS AVAILABLE TO BRING TO CLASSES? (GRADE 4)
GRADES/ASSESSMENTS:       N04, S04
GROUP LABEL:              COMP_BRG          LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                  C035703           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  3
001   CMPLAB-Y   (1         )  00           BRING COMPUTERS TO CLASS:  YES
002   CMPLAB-N   (2         )  10           BRING COMPUTERS TO CLASS:  NO
003   CMPLAB-?   (M         )  01           BRING COMPUTERS TO CLASS:  MISSING

CONDITIONING ID:          SCHL0083
DESCRIPTION:              WHAT PERCENT OF STUDENTS RECEIVE BILINGUAL EDUCATION?
GRADES/ASSESSMENTS:       N04, S04, N08, S08, N12
GROUP LABEL:              %BIL_ED           LENGTH OF CONTRAST FIELD       :  8
NAEP ID:                  C032004           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  9
001   %BIL_ED1   (1         )  00000000     PERCENT BILINGUAL ED:  NONE
002   %BIL_ED2   (2         )  10000000     PERCENT BILINGUAL ED:  1-5%
003   %BIL_ED3   (3         )  01000000     PERCENT BILINGUAL ED:  6-10%
004   %BIL_ED4   (4         )  00100000     PERCENT BILINGUAL ED:  11-25%
005   %BIL_ED5   (5         )  00010000     PERCENT BILINGUAL ED:  26-50%
006   %BIL_ED6   (6         )  00001000     PERCENT BILINGUAL ED:  51-75%
007   %BIL_ED7   (7         )  00000100     PERCENT BILINGUAL ED:  76-90%
```

Table F-4 (continued)
Conditioning Variables for Writing Main Samples

```
008   %BIL_ED8   (8          )  00000010     PERCENT BILINGUAL ED:  90-100%
009   %BIL_ED?   (M          )  00000001     PERCENT BILINGUAL ED:  MISSING

CONDITIONING ID:            SCHL0084
DESCRIPTION:                WHAT PERCENT OF STUDENTS RECEIVE ESL INSTRUCTION?
GRADES/ASSESSMENTS:         N04, S04, N08, S08, N12
GROUP LABEL:                %ESL_INS             LENGTH OF CONTRAST FIELD       :  8
NAEP ID:                    C032005              DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS                NUMBER OF SPECIFICATION RECORDS:  9
001   %ESL_IN1   (1          )  00000000     PERCENT ESL INSTRUCTION:  NONE
002   %ESL_IN2   (2          )  10000000     PERCENT ESL INSTRUCTION:  1-5%
003   %ESL_IN3   (3          )  01000000     PERCENT ESL INSTRUCTION:  6-10%
004   %ESL_IN4   (4          )  00100000     PERCENT ESL INSTRUCTION:  11-25%
005   %ESL_IN5   (5          )  00010000     PERCENT ESL INSTRUCTION:  26-50%
006   %ESL_IN6   (6          )  00001000     PERCENT ESL INSTRUCTION:  51-75%
007   %ESL_IN7   (7          )  00000100     PERCENT ESL INSTRUCTION:  76-90%
008   %ESL_IN8   (8          )  00000010     PERCENT ESL INSTRUCTION:  90-100%
009   %ESL_IN?   (M          )  00000001     PERCENT ESL INSTRUCTION:  MISSING

CONDITIONING ID:            SCHL0088
DESCRIPTION:                DOES SCHOOL ENCOURAGE PARENTS TO VISIT CLASSROOMS?
GRADES/ASSESSMENTS:         N04, S04, N08, S08, N12
GROUP LABEL:                PAR_VIS              LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                    C032208              DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS                NUMBER OF SPECIFICATION RECORDS:  4
001   PARVIS-R   (1          )  000          PARENTS VISIT CLASSROOMS:  ROUTINELY
002   PARVIS-O   (2          )  100          PARENTS VISIT CLASSROOMS:  OCCASIONALLY
003   PARVIS-N   (3          )  010          PARENTS VISIT CLASSROOMS:  NO
004   PARVIS-?   (M          )  001          PARENTS VISIT CLASSROOMS:  MISSING

CONDITIONING ID:            SCHL0089
DESCRIPTION:                DOES SCHOOL HAVE MINIMUM REQUIREMENT FOR HOMEWORK?
GRADES/ASSESSMENTS:         N04, S04, N08, S08, N12
GROUP LABEL:                MINREQHW             LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                    C032301              DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS                NUMBER OF SPECIFICATION RECORDS:  3
001   REQ_HW-Y   (1          )  00           MINIMUM HOMEWORK REQUIREMENT:  YES
002   REQ_HW-N   (2          )  10           MINIMUM HOMEWORK REQUIREMENT:  NO
003   REQ_HW-?   (M          )  01           MINIMUM HOMEWORK REQUIREMENT:  MISSING


CONDITIONING ID:            TCHR0002
DESCRIPTION:                TEACHER MATCH STATUS WITH STUDENT
GRADES/ASSESSMENTS:         N08
GROUP LABEL:                T_MATCH              LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                    TCHMTCH              DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS                NUMBER OF SPECIFICATION RECORDS:  3
001   TMCH-NO    (1,M        )  00           TEACHER MATCH:  NO MATCH
002   TMCH-PAR   (2          )  10           TEACHER MATCH:  PARTIAL MATCH
003   TMCH-COM   (3          )  01           TEACHER MATCH:  COMPLETE MATCH

CONDITIONING ID:            TCHR0003
DESCRIPTION:                TEACHER GENDER
GRADES/ASSESSMENTS:         N08
GROUP LABEL:                T_GENDER             LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                    T040001              DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS                NUMBER OF SPECIFICATION RECORDS:  3
001   T_MALE     (1          )  00           TEACHER GENDER:  MALE
002   T_FEMALE   (2          )  10           TEACHER GENDER:  FEMALE
003   T_SEX-?    (M,DNA      )  01           TEACHER GENDER:  MISSING OR DOES NOT APPLY

CONDITIONING ID:            TCHR0004
DESCRIPTION:                TEACHER RACE/ETHNICITY
GRADES/ASSESSMENTS:         N08
GROUP LABEL:                T_RACE               LENGTH OF CONTRAST FIELD       :  5
```

BEST COPY AVAILABLE

Table F-4 (continued)
Conditioning Variables for Writing Main Samples

```
NAEP ID:                  T040101          DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS           NUMBER OF SPECIFICATION RECORDS:  6
001   T_WHITE    (1         )  00000       TEACHER ETHNICITY:  WHITE
002   T_BLACK    (2         )  10000       TEACHER ETHNICITY:  BLACK
003   T_HISP     (3         )  01000       TEACHER ETHNICITY:  HISPANIC
004   T_ASIAN    (4         )  00100       TEACHER ETHNICITY:  ASIAN, PACIFIC ISLANDER
005   T_AM.IND   (5         )  00010       TEACHER ETHNICITY:  AMERICAN INDIAN, ALASKAN NATIVE
006   T_RACE-?   (M,DNA     )  00001       TEACHER ETHNICITY:  MISSING OR DOES NOT APPLY

CONDITIONING ID:           TCHR0005
DESCRIPTION:               TEACHER HISPANIC BACKGROUND
GRADES/ASSESSMENTS:       N08
GROUP LABEL:              T_HISPBK         LENGTH OF CONTRAST FIELD      :   5
NAEP ID:                  T040201          DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS           NUMBER OF SPECIFICATION RECORDS:  6
001   T_NONHSP   (1         )  00000       TEACHER HISPANIC BACKGROUND:  NOT HISPANIC
002   T_MEXICN   (2         )  10000       TEACHER HISPANIC BACKGROUND:  MEXICAN/MEXICAN AMERICAN
003   T_PUERTO   (3         )  01000       TEACHER HISPANIC BACKGROUND:  PUERTO RICAN
004   T_CUBAN    (4         )  00100       TEACHER HISPANIC BACKGROUND:  CUBAN
005   T_OTHER    (5         )  00010       TEACHER HISPANIC BACKGROUND:  OTHER
006   T_HISP-?   (M,DNA     )  00001       TEACHER HISPANIC BACKGROUND:  MISSING OR DOES NOT APPLY

CONDITIONING ID:           TCHR0006
DESCRIPTION:               YEARS TEACHING ELEMENTARY/SECONDARY SCHOOL
GRADES/ASSESSMENTS:       N08
GROUP LABEL:              T_YRSEXP         LENGTH OF CONTRAST FIELD      :   5
NAEP ID:                  T040301          DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS           NUMBER OF SPECIFICATION RECORDS:  6
001   T_YREXP1   (1         )  00000       YEARS TEACHING:  2 OR LESS YEARS
002   T_YREXP2   (2         )  10000       YEARS TEACHING:  3-5 YEARS
003   T_YREXP3   (3         )  01000       YEARS TEACHING:  6-10 YEARS
004   T_YREXP4   (4         )  00100       YEARS TEACHING:  11-24 YEARS
005   T_YREXP5   (5         )  00010       YEARS TEACHING:  25 OR MORE YEARS
006   T_YREXP?   (M,DNA     )  00001       YEARS TEACHING:  MISSING OR DOES NOT APPLY

CONDITIONING ID:           TCHR0007
DESCRIPTION:               TYPE OF TEACHING CERTIFICATION
GRADES/ASSESSMENTS:       N08
GROUP LABEL:              TCH_CERT         LENGTH OF CONTRAST FIELD      :   3
NAEP ID:                  T040401          DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS           NUMBER OF SPECIFICATION RECORDS:  4
001   TCERT-NO   (1         )  000         TEACHING CERTIFICATION:  NONE, TEMPORARY, PROVISIONAL
002   TCERT-RG   (2         )  100         TEACHING CERTIFICATION:  REGULAR, NOT HIGHEST AVAILABLE
003   TCERT-HI   (3         )  010         TEACHING CERTIFICATION:  HIGHEST AVAILABLE
004   TCERT-?    (M,DNA     )  001         TEACHING CERTIFICATION:  MISSING OR DOES NOT APPLY

CONDITIONING ID:           TCHR0008
DESCRIPTION:               TEACHER GENERAL CERTIFICATION (ELEMENTARY, MIDDLE/JUNIOR HS EDUCATION)
GRADES/ASSESSMENTS:       N08
GROUP LABEL:              CERT-GEN         LENGTH OF CONTRAST FIELD      :   3
NAEP ID:                  T040501          DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS           NUMBER OF SPECIFICATION RECORDS:  4
001   CERTG-Y    (1         )  000         GENERAL CERTIFICATION:  YES
002   CERTG-N    (2         )  100         GENERAL CERTIFICATION:  NO
003   CERTG-NS   (3         )  010         GENERAL CERTIFICATION:  NOT OFFERED IN STATE
004   CERTG-?    (M,DNA     )  001         GENERAL CERTIFICATION:  MISSING OR DOES NOT APPLY

CONDITIONING ID:           TCHR0009
DESCRIPTION:               TEACHER'S HIGHEST ACADEMIC DEGREE
GRADES/ASSESSMENTS:       N08
GROUP LABEL:              T_DEGREE         LENGTH OF CONTRAST FIELD      :   6
NAEP ID:                  T040601          DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS           NUMBER OF SPECIFICATION RECORDS:  7
001   <BACHLRS   (1         )  000000      TEACHER DEGREE:  LESS THAN A BACHELOR'S DEGREE
002   BACHELRS   (2         )  100000      TEACHER DEGREE:  BACHELOR'S DEGREE
```

Table F-4 (continued)
Conditioning Variables for Writing Main Samples

```
003   MASTERS    (3              )  010000       TEACHER DEGREE:  MASTER'S DEGREE
004   SPECLIST   (4              )  001000       TEACHER DEGREE:  EDUCATION SPECIALIST
005   DOCTORAT   (5              )  000100       TEACHER DEGREE:  DOCTORATE
006   PROFESSL   (6              )  000010       TEACHER DEGREE:  PROFESSIONAL DEGREE
007   DEGREE-?   (M,DNA          )  000001       TEACHER DEGREE:  MISSING OR DOES NOT APPLY

CONDITIONING ID:            TCHR0010
DESCRIPTION:                TEACHER UNDERGRADUATE MAJOR IN EDUCATION GRADES/ASSESSMENTS:        N08
GROUP LABEL:                UGRAD_ED            LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                    T040701             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   2
001   UGR_ED-?   (M,DNA          )  0            UNDERGRADUATE EDUCATION MAJOR:  MISSING OR DOES NOT APPLY
002   UGR_ED-Y   (1              )  1            UNDERGRADUATE EDUCATION MAJOR:  YES

CONDITIONING ID:            TCHR0011
DESCRIPTION:                TEACHER UNDERGRADUATE MAJOR IN ENGLISH
GRADES/ASSESSMENTS:         N08
GROUP LABEL:                UGRAD_EG            LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                    T040706             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   2
001   UGR_EG-?   (M,DNA          )  0            UNDERGRADUATE ENGLISH MAJOR:  MISSING OR DOES NOT APPLY
002   UGR_EG-Y   (1              )  1            UNDERGRADUATE ENGLISH MAJOR:  YES

CONDITIONING ID:            TCHR0012
DESCRIPTION:                TEACHER GRADUATE MAJOR IN EDUCATION
GRADES/ASSESSMENTS:         N08
GROUP LABEL:                GRAD_ED             LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                    T040801             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   2
001   GR_ED-?    (M,DNA          )  0            GRADUATE EDUCATION MAJOR:  MISSING OR DOES NOT APPLY
002   GR_ED-Y    (1              )  1            GRADUATE EDUCATION MAJOR:  YES

CONDITIONING ID:            TCHR0013
DESCRIPTION:                TEACHER GRADUATE MAJOR IN ENGLISH
GRADES/ASSESSMENTS:         N08
GROUP LABEL:                GRAD_EG             LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                    T040807             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   2
001   GR_EG-?    (M,DNA          )  0            GRADUATE ENGLISH MAJOR:  MISSING OR DOES NOT APPLY
002   GR_EG-Y    (1              )  1            GRADUATE ENGLISH MAJOR:  YES

CONDITIONING ID:            TCHR0014
DESCRIPTION:                NO TEACHER GRADUATE-LEVEL STUDY
GRADES/ASSESSMENTS:         N08
GROUP LABEL:                GRAD_NO             LENGTH OF CONTRAST FIELD       :   1
NAEP ID:                    T040806             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   2
001   GR_NO-?    (M,DNA          )  0            NO GRAUDATE STUDY:  MISSING OR DOES NOT APPLY
002   GR_NO-Y    (1              )  1            NO GRADUATE STUDY:  YES

CONDITIONING ID:            TCHR0017
DESCRIPTION:                NUMBER OF STUDENTS IN CLASS
GRADES/ASSESSMENTS:         N08
GROUP LABEL:                T_NCLASS            LENGTH OF CONTRAST FIELD       :   5
NAEP ID:                    TCHNCLS             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   6
001   T_NCLAS1   (0-20           )  00000        CLASS SIZE:  0-20
002   T_NCLAS2   (21-25          )  10000        CLASS SIZE:  21-25
003   T_NCLAS3   (26-30          )  01000        CLASS SIZE:  26-30
004   T_NCLAS4   (31-35          )  00100        CLASS SIZE:  31-35
005   T_NCLAS5   (36-61          )  00010        CLASS SIZE:  36-60
006   T_NCLAS?   (M              )  00001        CLASS SIZE:  MISSING

CONDITIONING ID:            TCHR0026
DESCRIPTION:                TEACHER CONTROL IN SELECTING INSTRUCTIONAL MATERIALS
```

Table F-4 (continued)
Conditioning Variables for Writing Main Samples

```
GRADES/ASSESSMENTS:          N08
GROUP LABEL:                 CNTL_IM        LENGTH OF CONTRAST FIELD        :  5
NAEP ID:                     T041101        DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:            CLASS          NUMBER OF SPECIFICATION RECORDS:   6
001   CNTL_IM1   (1            )  00000     CONTROL IN INSTRUCTIONAL MATERIALS:  COMPLETE
002   CNTL_IM2   (2            )  10000     CONTROL IN INSTRUCTIONAL MATERIALS:  A LOT
003   CNTL_IM3   (3            )  01000     CONTROL IN INSTRUCTIONAL MATERIALS:  SOME
004   CNTL_IM4   (4            )  00100     CONTROL IN INSTRUCTIONAL MATERIALS:  VERY LITTLE
005   CNTL_IM5   (5            )  00010     CONTROL IN INSTRUCTIONAL MATERIALS:  NONE
006   CNTL_IM?   (M,DNA        )  00001     CONTROL IN INSTRUCTIONAL MATERIALS:  MISS OR DOES NOT APPLY

CONDITIONING ID:             TCHR0027
DESCRIPTION:                 TEACHER CONTROL IN DECIDING COURSE CONTENT
GRADES/ASSESSMENTS:          N08
GROUP LABEL:                 CNTL_CC        LENGTH OF CONTRAST FIELD        :  5
NAEP ID:                     T041102        DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:            CLASS          NUMBER OF SPECIFICATION RECORDS:   6
001   CNTL_CC1   (1            )  00000     CONTROL IN COURSE CONTENT:  COMPLETE
002   CNTL_CC2   (2            )  10000     CONTROL IN COURSE CONTENT:  A LOT
003   CNTL_CC3   (3            )  01000     CONTROL IN COURSE CONTENT:  SOME
004   CNTL_CC4   (4            )  00100     CONTROL IN COURSE CONTENT:  VERY LITTLE
005   CNTL_CC5   (5            )  00010     CONTROL IN COURSE CONTENT:  NONE
006   CNTL_CC?   (M,DNA        )  00001     CONTROL IN COURSE CONTENT:  MISSING OR DOES NOT APPLY

CONDITIONING ID:             TCHR0028
DESCRIPTION:                 TEACHER CONTROL IN DECIDING SEQUENCE OF CONTENT
GRADES/ASSESSMENTS:          N08
GROUP LABEL:                 CNTL_CS        LENGTH OF CONTRAST FIELD        :  5
NAEP ID:                     T041103        DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:            CLASS          NUMBER OF SPECIFICATION RECORDS:   6
001   CNTL_CS1   (1            )  00000     CONTROL IN CONTENT SEQUENCE:  COMPLETE
002   CNTL_CS2   (2            )  10000     CONTROL IN CONTENT SEQUENCE:  A LOT
003   CNTL_CS3   (3            )  01000     CONTROL IN CONTENT SEQUENCE:  SOME
004   CNTL_CS4   (4            )  00100     CONTROL IN CONTENT SEQUENCE:  VERY LITTLE
005   CNTL_CS5   (5            )  00010     CONTROL IN CONTENT SEQUENCE:  NONE
006   CNTL_CS?   (M,DNA        )  00001     CONTROL IN CONTENT SEQUENCE:  MISSING OR DOES NOT APPLY

CONDITIONING ID:             TCHR0030
DESCRIPTION:                 DO YOU HAVE CERTIFICATION IN JR HIGH/SECONDARY ENGLISH
GRADES/ASSESSMENTS:          N08
GROUP LABEL:                 CERT_JHE       LENGTH OF CONTRAST FIELD        :  2
NAEP ID:                     T040503        DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:            CLASS          NUMBER OF SPECIFICATION RECORDS:   3
001   JHSENG-Y   (1            )  00        CERTIFICATION IN JR. HIGH ENGLISH:  YES
002   JHSENG-N   (2,3          )  10        CERTIFICATION IN JR. HIGH ENGLISH:  NO OR NOT OFFERED
003   JHSENG-?   (M,DNA        )  01        CERTIFICATION IN JR. HIGH ENGLISH:  MISS OR DOES NOT APPLY

CONDITIONING ID:             TCHR0031
DESCRIPTION:                 DO YOU HAVE ANY OTHER TYPE OF CERTIFICATION
GRADES/ASSESSMENTS:          N08
GROUP LABEL:                 CERT_JHO       LENGTH OF CONTRAST FIELD        :  2
NAEP ID:                     T040505        DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:            CLASS          NUMBER OF SPECIFICATION RECORDS:   3
001   CRTOTH-Y   (1            )  00        CERTIFICATION IN OTHER AREAS:  YES
002   CRTOTH-N   (2,3          )  10        CERTIFICATION IN OTHER AREAS:  NO, OR NOT OFFERED IN STATE
003   CRTOTH-?   (M,DNA        )  01        CERTIFICATION IN OTHER AREAS:  MISSING OR DOES NOT APPLY

CONDITIONING ID:             TCHR0032
DESCRIPTION:                 TEACHER UNDERGRADUATE MAJOR IN ENGLISH EDUCATION
GRADES/ASSESSMENTS:          N08
GROUP LABEL:                 UGRAD_EE       LENGTH OF CONTRAST FIELD        :  1
NAEP ID:                     T040707        DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:            CLASS          NUMBER OF SPECIFICATION RECORDS:   2
001   UG_EGE-?   (M,DNA        )  0         UNDERGRADUATE ENGLISH EDUCATION MAJOR:  MISSING OR NOT APPLY
002   UG_EGE-Y   (1            )  1         UNDERGRADUATE ENGLISH EDUCATION MAJOR:  YES
```

Table F-4 (continued)
Conditioning Variables for Writing Main Samples

```
CONDITIONING ID:          TCHR0033
DESCRIPTION:              TEACHER UNDERGRADUATE MAJOR IN OTHER FIELDS
GRADES/ASSESSMENTS:       N08
GROUP LABEL:              UGRAD_OT          LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                  T040705           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  2
001   UG_OTH-?   (M,DNA      )  0           UNDERGRADUATE MAJOR IN OTHER FIELDS:  MISSING OR NOT APPLY
002   UG_OTH-Y   (1          )  1           UNDERGRADUATE MAJOR IN OTHER FIELDS:  YES

CONDITIONING ID:          TCHR0034
DESCRIPTION:              TEACHER GRADUATE MAJOR IN OTHER FIELDS
GRADES/ASSESSMENTS:       N08
GROUP LABEL:              GRAD_OTH          LENGTH OF CONTRAST FIELD       :  1
NAEP ID:                  T040805           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  2
001   GR_OTH-?   (M,DNA      )  0           GRADUATE MAJOR IN OTHER FIELDS:  MISSING OR DOES NOT APPLY
002   GR_OTH-Y   (1          )  1           GRADUATE MAJOR IN OTHER FIELDS:  YES

CONDITIONING ID:          TCHR0035
DESCRIPTION:              NO SPECIFIC TRAINING FOR TEACHING WRITING
GRADES/ASSESSMENTS:       N08
GROUP LABEL:              WRT_TRNG          LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                  T041901           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  3
001   WRT_TR-?   (DNA        )  00          NO SPECIFIC TRAINING FOR TEACHING WRITING: DOES NOT APPLY
002   WRT_TR-O   (M          )  10          NO SPECIFIC TRAINING FOR TEACHING WRITING: OMIT--ASSUMED YES
003   WRT_TR-N   (1          )  01          NO SPECIFIC TRAINING FOR TEACHING WRITING: NO

CONDITIONING ID:          TCHR0036
DESCRIPTION:              HOW MANY WRITING COURSES DID YOU TAKE?
GRADES/ASSESSMENTS:       N08
GROUP LABEL:              WRT_COUR          LENGTH OF CONTRAST FIELD       :  5
NAEP ID:                  T041801           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  6
001   WRT_CR-0   (1          )  00000       WRITING COURSES TAKEN: NONE
002   WRT_CR-1   (2          )  10000       WRITING COURSES TAKEN: ONE
003   WRT_CR-2   (3          )  01000       WRITING COURSES TAKEN: TWO
004   WRT_CR-3   (4          )  00100       WRITING COURSES TAKEN: THREE OR FOUR
005   WRT_CR-4   (5          )  00010       WRITING COURSES TAKEN: FIVE OR MORE
006   WRT_CR-?   (M,DNA      )  00001       WRITING COURSES TAKEN: MISSING OR DOES NOT APPLY

CONDITIONING ID:          TCHR0037
DESCRIPTION:              HOW MUCH TIME SPENT ATTENDING CONFERENCES
GRADES/ASSESSMENTS:       N08
GROUP LABEL:              CONF_TIM          LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                  T042401           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  5
001   CONF___0   (1          )  0000        TIME SPENT ATTENDING CONFERENCES: NONE
002   CONF__<6   (2          )  1000        TIME SPENT ATTENDING CONFERENCES: LESS THAN 6 HOURS
003   CON_6-15   (3          )  0100        TIME SPENT ATTENDING CONFERENCES: 6 TO 15 HOURS
004   CONF_16+   (4,5        )  0010        TIME SPENT ATTENDING CONFERENCES: MORE THAN 16 HOURS
005   CONF___?   (M,DNA      )  0001        TIME SPENT ATTENDING CONFERENCES: MISSING OR DOES NOT APPLY

CONDITIONING ID:          TCHR0038
DESCRIPTION:              HOW MUCH TIME SPENT READING PUBLICATIONS GRADES/ASSESSMENTS:        N08
GROUP LABEL:              READ_PUB          LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                  T042402           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:         CLASS             NUMBER OF SPECIFICATION RECORDS:  5
001   RDPB___0   (1          )  0000        TIME SPENT READING PUBLICATIONS: NONE
002   RDPB__<6   (2          )  1000        TIME SPENT READING PUBLICATIONS: LESS THAN 6 HOURS
003   RDP_6-15   (3          )  0100        TIME SPENT READING PUBLICATIONS: 6 TO 15 HOURS
004   RDPB_16+   (4,5        )  0010        TIME SPENT READING PUBLICATIONS: MORE THAN 16 HOURS
005   RDPB___?   (M,DNA      )  0001        TIME SPENT READING PUBLICATIONS: MISSING OR DOES NOT APPLY
```

Table F-4 (continued)
Conditioning Variables for Writing Main Samples

```
CONDITIONING ID:          TWRI0001
DESCRIPTION:              ACCORDING TO TEACHER ARE STUDENTS ASSIGNED TO CLASSES BY ABILITY
GRADES/ASSESSMENTS:       N08
GROUP LABEL:              ST_BY_AB            LENGTH OF CONTRAST FIELD       :   2
NAEP ID:                  T048101             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:         CLASS               NUMBER OF SPECIFICATION RECORDS:   3
001   ST_AB__Y   (1          )  00            STUDENTS ASSIGNED BY ABILITY: YES
002   ST_AB__N   (2          )  10            STUDENTS ASSIGNED BY ABILITY: NO
003   ST_AB__?   (M,DNA      )  01            STUDENTS ASSIGNED BY ABILITY: MISSING OR DOES NOT APPLY

CONDITIONING ID:          TWRI0002
DESCRIPTION:              WHAT IS WRITING ABILITY OF STUDENTS' IN THIS CLASS?
GRADES/ASSESSMENTS:       N08
GROUP LABEL:              SWRT_ABL            LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                  T048201             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:         CLASS               NUMBER OF SPECIFICATION RECORDS:   5
001   SWTAB__H   (1          )  0000          STUDENTS' WRITING ABILITY: MOSTLY HIGH
002   ST_AB__A   (2          )  1000          STUDENTS' WRITING ABILITY: MOSTLY AVERAGE
003   ST_AB__L   (3          )  0100          STUDENTS' WRITING ABILITY: MOSTLY LOW
004   ST_AB__M   (4          )  0010          STUDENTS' WRITING ABILITY: MIXED
005   ST_AB__?   (M,DNA      )  0001          STUDENTS' WRITING ABILITY: MISSING OR DOES NOT APPLY

CONDITIONING ID:          TWRI0003
DESCRIPTION:              ACCORDING TO TEACHER CREATE GROUPS FOR INSTRUCTION ON BASIS OF ABILITY
GRADES/ASSESSMENTS:       N08
GROUP LABEL:              GR_BY_AB            LENGTH OF CONTRAST FIELD       :   2
NAEP ID:                  T048301             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:         CLASS               NUMBER OF SPECIFICATION RECORDS:   3
001   GR_AB__Y   (1          )  00            GROUPS ASSIGNED BY ABILITY: YES
002   GR_AB__N   (2          )  10            GROUPS ASSIGNED BY ABILITY: NO
003   GR_AB__?   (M,DNA      )  01            GROUPS ASSIGNED BY ABILITY: MISSING OR DOES NOT APPLY

CONDITIONING ID:          TWRI0004
DESCRIPTION:              TIME SPENT ON WRITING OUTSIDE OF CLASS PER WEEK
GRADES/ASSESSMENTS:       N08
GROUP LABEL:              TWRT_OUT            LENGTH OF CONTRAST FIELD       :   5
NAEP ID:                  T048401             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:         CLASS               NUMBER OF SPECIFICATION RECORDS:   6
001   TWOUT__0   (1          )  00000         TIME SPENT ON WRITING OUTSIDE OF CLASS: NONE
002   TWOUT_<1   (2          )  10000         TIME SPENT ON WRITING OUTSIDE OF CLASS: < ONE HOUR
003   TWOUT__1   (3          )  01000         TIME SPENT ON WRITING OUTSIDE OF CLASS: ONE HOUR
004   TWOUT__2   (4          )  00100         TIME SPENT ON WRITING OUTSIDE OF CLASS: 2 HOURS
005   TWOUT_3+   (5          )  00010         TIME SPENT ON WRITING OUTSIDE OF CLASS: 3 OR MORE
006   TWOUT__?   (M,DNA      )  00001         TIME SPENT ON WRITING OUTSIDE OF CLASS: MISSING OR NOT APPLY

CONDITIONING ID:          TWRI0005
DESCRIPTION:              TIME SPENT HELPING STUDENTS WITH WRITING PER WEEK
GRADES/ASSESSMENTS:       N08
GROUP LABEL:              HLPST WR            LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                  T048501             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:         CLASS               NUMBER OF SPECIFICATION RECORDS:   5
001   HSTW_<30   (1          )  0000          TIME SPENT HELPING WITH WRITING PER WEEK: 30 MINUTES OR LESS
002   HSTW__60   (2          )  1000          TIME SPENT HELPING WITH WRITING PER WEEK: 60 MINUTES
003   HSTW__90   (3          )  0100          TIME SPENT HELPING WITH WRITING PER WEEK: 90 MINUTES
004   HSTW_120   (4          )  0010          TIME SPENT HELPING WITH WRITING PER WEEK: TWO OR MORE HOURS
005   HSTW___?   (M,DNA      )  0001          TIME SPENT HELPING WITH WRITING PER WEEK: MISS, OR NOT APPLY

CONDITIONING ID:          TWRI0006
DESCRIPTION:              DO YOU USE WORKBOOKS OR WORKSHEETS FOR INSTRUCTION?
GRADES/ASSESSMENTS:       N08
GROUP LABEL:              WBK_INST            LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                  T048601             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:         CLASS               NUMBER OF SPECIFICATION RECORDS:   4
001   WBK__CEN   (1          )  000           WORKBOOKS USED FOR INSTRUCTION: YES AS A CENTRAL PART
002   WBK__SUP   (2          )  100           WORKBOOKS USED FOR INSTRUCTION: YES AS A SUPPLEMENTAL
```

Table F-4 (continued)
Conditioning Variables for Writing Main Samples

```
003   WBK___NO    (3           )  010          WORKBOOKS USED FOR INSTRUCTION: NO
004   WBK____?    (M,DNA       )  001          WORKBOOKS USED FOR INSTRUCTION: MISSING OR DOES NOT APPLY

CONDITIONING ID:               TWRI0007
DESCRIPTION:                   DO YOU USE TEXTBOOKS FOR INSTRUCTION?
GRADES/ASSESSMENTS:            N08
GROUP LABEL:                   TXT_INST               LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                       T048701                DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:              CLASS                  NUMBER OF SPECIFICATION RECORDS:  4
001   TXT__CEN    (1           )  000          TEXTBOOKS USED FOR INSTRUCTION: YES AS A CENTRAL PART
002   TXT__SUP    (2           )  100          TEXTBOOKS USED FOR INSTRUCTION: YES AS A SUPPLEMENTAL
003   TXT___NO    (3           )  010          TEXTBOOKS USED FOR INSTRUCTION: NO
004   TXT____?    (M,DNA       )  001          TEXTBOOKS USED FOR INSTRUCTION: MISSING OR DOES NOT APPLY

CONDITIONING ID:               TWRI0008
DESCRIPTION:                   INSTRUCTIONAL APPROACH GRAMMAR AND SKILLS-BASED
GRADES/ASSESSMENTS:            N08
GROUP LABEL:                   GRAMINST               LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                       T048801                DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:              CLASS                  NUMBER OF SPECIFICATION RECORDS:  4
001   GRMI_CEN    (1           )  000          INSTRUCT APPROACH GRAMMAR/SKILLS-BASED: YES CENTRAL PART
002   GRMI_SUP    (2           )  100          INSTRUCT APPROACH GRAMMAR/SKILLS-BASED: YES SUPPLEMENTAL
003   GRMI__NO    (3           )  010          INSTRUCT APPROACH GRAMMAR/SKILLS-BASED: NO
004   GRMI___?    (M,DNA       )  001          INSTRUCT APPROACH GRAMMAR/SKILLS-BASED: MISSING OR NOT APPLY

CONDITIONING ID:               TWRI0009
DESCRIPTION:                   INSTRUCTIONAL APPROACH WRITING PROCESS
GRADES/ASSESSMENTS:            N08
GROUP LABEL:                   PROCINST               LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                       T048802                DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:              CLASS                  NUMBER OF SPECIFICATION RECORDS:  4
001   PROC_CEN    (1           )  000          INSTRUCTIONAL APPROACH WRITING PROCESS: YES CENTRAL PART
002   PROC_SUP    (2           )  100          INSTRUCTIONAL APPROACH WRITING PROCESS: YES SUPPLEMENTAL
003   PROC__NO    (3           )  010          INSTRUCTIONAL APPROACH WRITING PROCESS: NO
004   PROC___?    (M,DNA       )  001          INSTRUCTIONAL APPROACH WRITING PROCESS: MISSING OR NOT APPLY

CONDITIONING ID:               TWRI0010
DESCRIPTION:                   INSTRUCTIONAL APPROACH INTEGRATING READING AND WRITING
GRADES/ASSESSMENTS:            N08
GROUP LABEL:                   R+W_INST               LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                       T048803                DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:              CLASS                  NUMBER OF SPECIFICATION RECORDS:  4
001   R+WI_CEN    (1           )  000          INSTRUCT APPROACH INTEGRATING RDG/WRTNG: YES CENTRAL PART
002   R+WI_SUP    (2           )  100          INSTRUCT APPROACH INTEGRATING RDG/WRTNG: YES SUPPLEMENT
003   R+WI__NO    (3           )  010          INSTRUCT APPROACH INTEGRATING RDG/WRTNG: NO
004   R+WI___?    (M,DNA       )  001          INSTRUCT APPROACH INTEGRATING RDG/WRTNG: MISSNG OR NOT APPLY

CONDITIONING ID:               TWRI0011
DESCRIPTION:                   INSTRUCTIONAL APPROACH WRITING ABOUT LITERATURE
GRADES/ASSESSMENTS:            N08
GROUP LABEL:                   LIT_INST               LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                       T048804                DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:              CLASS                  NUMBER OF SPECIFICATION RECORDS:  4
001   LITI_CEN    (1           )  000          INSTRUCTIONAL APPROACH WRITING ABOUT LIT: YES CENTRAL PART
002   LITI_SUP    (2           )  100          INSTRUCTIONAL APPROACH WRITING ABOUT LIT: YES SUPPLEMENTAL
003   LITI__NO    (3           )  010          INSTRUCTIONAL APPROACH WRITING ABOUT LIT: NO
004   LITI___?    (M,DNA       )  001          INSTRUCTNL APPROACH WRITING ABOUT LIT: MISSING OR NOT APPLY

CONDITIONING ID:               TWRI0012
DESCRIPTION:                   INSTRUCTIONAL APPROACH WRITING ACROSS SUBJECTS
GRADES/ASSESSMENTS:            N08
GROUP LABEL:                   WRT_ACSB               LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                       T048805                DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:              CLASS                  NUMBER OF SPECIFICATION RECORDS:  4
001   ACSB_CEN    (1           )  000          INSTRUCT APPROACH WRITING ACROSS SUBJECTS: YES CENTRAL PART
```

Table F-4 (continued)
Conditioning Variables for Writing Main Samples

```
002   ACSB_SUP   (2          )  100        INSTRUCT APPROACH WRITING ACROSS SUBJECTS: YES SUPPLEMENTAL
003   ACSB__NO   (3          )  010        INSTRUCT APPROACH WRITING ACROSS SUBJECTS: NO
004   ACSB___?   (M,DNA      )  001        INSTRUCT APPROACH WRITING ACROSS SUBJ: MISSING OR NOT APPLY

CONDITIONING ID:            TWRI0013
DESCRIPTION:                HOW OFTEN DO YOU ASSIGN PAPERS OF 1-2 PARAGRAPHS
GRADES/ASSESSMENTS:         W08
GROUP LABEL:                T1-2PARA            LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                    T048901             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:  5
001   TPAR_DAY   (1          )  0000       HOW OFTEN ASSIGN 1-2 PARAGRAPHS: ALMOST EVERY DAY
002   TPAR__WK   (2          )  1000       HOW OFTEN ASSIGN 1-2 PARAGRAPHS: ONCE OR TWICE A WEEK
003   TPAR_MON   (3          )  0100       HOW OFTEN ASSIGN 1-2 PARAGRAPHS: ONCE OR TWICE A MONTH
004   TPAR_NEV   (4          )  0010       HOW OFTEN ASSIGN 1-2 PARAGRAPHS: NEVER OR HARDLY EVER
005   TPAR___?   (M,DNA      )  0001       HOW OFTEN ASSIGN 1-2 PARAGRAPHS: MISSING OR DOES NOT APPL

CONDITIONING ID:            TWRI0014
DESCRIPTION:                HOW OFTEN DO YOU ASSIGN PAPERS OF 1-2 PAGES
GRADES/ASSESSMENTS:         W08
GROUP LABEL:                T1-2PPS             LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                    T048902             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:  5
001   TPPS_DAY   (1          )  0000       HOW OFTEN ASSIGN PAPERS OF 1-2 PAGES: ALMOST EVERY DAY
002   TPPS__WK   (2          )  1000       HOW OFTEN ASSIGN PAPERS OF 1-2 PAGES: ONCE OR TWICE A WEEK
003   TPPS_MON   (3          )  0100       HOW OFTEN ASSIGN PAPERS OF 1-2 PAGES: ONCE OR TWICE A MONTH
004   TPPS_NEV   (4          )  0010       HOW OFTEN ASSIGN PAPERS OF 1-2 PAGES: NEVER OR HARDLY EVER
005   TPPS___?   (M,DNA      )  0001       HOW OFTEN ASSIGN PAPERS OF 1-2 PAGES: MISSING OR NOT APPLY

CONDITIONING ID:            TWRI0015
DESCRIPTION:                HOW OFTEN DO YOU ASSIGN PAPERS OF 3+ PAGES
GRADES/ASSESSMENTS:         W08
GROUP LABEL:                T3+_PPS             LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                    T048903             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:  5
001   TPP3_DAY   (1          )  0000       HOW OFTEN ASSIGN PAPERS OF 3+ PAGES: ALMOST EVERY DAY
002   TPP3__WK   (2          )  1000       HOW OFTEN ASSIGN PAPERS OF 3+ PAGES: ONCE OR TWICE A WEEK
003   TPP3_MON   (3          )  0100       HOW OFTEN ASSIGN PAPERS OF 3+ PAGES: ONCE OR TWICE A MONTH
004   TPP3_NEV   (4          )  0010       HOW OFTEN ASSIGN PAPERS OF 3+ PAGES: NEVER OR HARDLY EVER
005   TPPS___?   (M,DNA      )  0001       HOW OFTEN ASSIGN PAPERS OF 3+ PAGES: MISSING OR NOT APPLY

CONDITIONING ID:            TWRI0016
DESCRIPTION:                HOW OFTEN DO YOU ASSIGN A REPORT OR SUMMARY
GRADES/ASSESSMENTS:         W08
GROUP LABEL:                ASG_REPT            LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                    T049001             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:  5
001   ASGR_DAY   (1          )  0000       HOW OFTEN ASSIGN A REPORT OR SUMMARY: ALMOST EVERY DAY
002   ASGR__WK   (2          )  1000       HOW OFTEN ASSIGN A REPORT OR SUMMARY: ONCE OR TWICE A WEEK
003   ASGR_MON   (3          )  0100       HOW OFTEN ASSIGN A REPORT OR SUMMARY: ONCE OR TWICE A MONTH
004   ASGR_NEV   (4          )  0010       HOW OFTEN ASSIGN A REPORT OR SUMMARY: NEVER OR HARDLY EVER
005   ASGREP_?   (M,DNA      )  0001       HOW OFTEN ASSIGN A REPORT OR SUMMARY: MISSING OR NOT APPLY

CONDITIONING ID:            TWRI0017
DESCRIPTION:                HOW OFTEN DO YOU ASSIGN AN ANALYTICAL ESSAY
GRADES/ASSESSMENTS:         W08
GROUP LABEL:                ASG_ESSY            LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                    T049002             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:  5
001   ASGE_DAY   (1          )  0000       HOW OFTEN ASSIGN AN ANALYTICAL ESSAY: ALMOST EVERY DAY
002   ASGE__WK   (2          )  1000       HOW OFTEN ASSIGN AN ANALYTICAL ESSAY: ONCE OR TWICE A WEEK
003   ASGE_MON   (3          )  0100       HOW OFTEN ASSIGN AN ANALYTICAL ESSAY: ONCE OR TWICE A MONTH
004   ASGE_NEV   (4          )  0010       HOW OFTEN ASSIGN AN ANALYTICAL ESSAY: NEVER OR HARDLY EVER
005   ASGESS_?   (M,DNA      )  0001       HOW OFTEN ASSIGN AN ANALYTICAL ESSAY: MISSING OR NOT APPLY

CONDITIONING ID:            TWRI0018
```

Table F-4 (continued)
Conditioning Variables for Writing Main Samples

```
DESCRIPTION:               HOW OFTEN DO YOU ASSIGN A PERSUASIVE ESSAY
GRADES/ASSESSMENTS:        N08
GROUP LABEL:               ASG_PERS          LENGTH OF CONTRAST FIELD        :  4
NAEP ID:                   T049003           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:   5
001   ASPE_DAY   (1          )  0000         HOW OFTEN ASSIGN A PERSUASIVE ESSAY: ALMOST EVERY DAY
002   ASPE__WK   (2          )  1000         HOW OFTEN ASSIGN A PERSUASIVE ESSAY: ONCE OR TWICE A WEEK
003   ASPE_MON   (3          )  0100         HOW OFTEN ASSIGN A PERSUASIVE ESSAY: ONCE OR TWICE A MONTH
004   ASPE_NEV   (4          )  0010         HOW OFTEN ASSIGN A PERSUASIVE ESSAY: NEVER OR HARDLY EVER
005   ASPESS_?   (M,DNA      )  0001         HOW OFTEN ASSIGN A PERSUASIVE ESSAY: MISSING OR NOT APPLY

CONDITIONING ID:           TWRI0019
DESCRIPTION:               HOW OFTEN DO YOU ASSIGN A NARRATIVE ESSAY OR STORY
GRADES/ASSESSMENTS:        N08
GROUP LABEL:               ASG_NARR          LENGTH OF CONTRAST FIELD        :  4
NAEP ID:                   T049004           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:   5
001   ASNE_DAY   (1          )  0000         HOW OFTEN ASSIGN A NARRATIVE ESSAY: ALMOST EVERY DAY
002   ASNE__WK   (2          )  1000         HOW OFTEN ASSIGN A NARRATIVE ESSAY: ONCE OR TWICE A WEEK
003   ASNE_MON   (3          )  0100         HOW OFTEN ASSIGN A NARRATIVE ESSAY: ONCE OR TWICE A MONTH
004   ASNE_NEV   (4          )  0010         HOW OFTEN ASSIGN A NARRATIVE ESSAY: NEVER OR HARDLY EVER
005   ASNESS_?   (M,DNA      )  0001         HOW OFTEN ASSIGN A NARRATIVE ESSAY: MISSING OR NOT APPLY

CONDITIONING ID:           TWRI0020
DESCRIPTION:               HOW OFTEN DO YOU ASSIGN WRITING IN A JOURNAL
GRADES/ASSESSMENTS:        N08
GROUP LABEL:               ASG_JOUR          LENGTH OF CONTRAST FIELD        :  4
NAEP ID:                   T049005           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:   5
001   ASJO_DAY   (1          )  0000         HOW OFTEN ASSIGN WRITING IN A JOURNAL: ALMOST EVERY DAY
002   ASJO__WK   (2          )  1000         HOW OFTEN ASSIGN WRITING IN A JOURNAL: ONCE OR TWICE A WEEK
003   ASJO_MON   (3          )  0100         HOW OFTEN ASSIGN WRITING IN A JOURNAL: ONCE OR TWICE A MONTH
004   ASJO_NEV   (4          )  0010         HOW OFTEN ASSIGN WRITING IN A JOURNAL: NEVER OR HARDLY EVER
005   ASJOUR_?   (M,DNA      )  0001         HOW OFTEN ASSIGN WRITING IN A JOURNAL: MISSING OR NOT APPLY

CONDITIONING ID:           TWRI0021
DESCRIPTION:               HOW IMPORTANT IN GRADING? SPELLING/PUNCTUATION/GRAMMAR
GRADES/ASSESSMENTS:        N08
GROUP LABEL:               IMP__SPG          LENGTH OF CONTRAST FIELD        :  3
NAEP ID:                   T049101           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:   4
001   IMPGMR_V   (1          )  000          HOW IMPORTANT IN GRADING? SPELL/PUNC/GRAMMAR: VERY IMPORTANT
002   IMPGMR_M   (2          )  100          HOW IMPORTANT IN GRADING? SPELL/PUNC/GRAMMAR: MODER IMPORT
003   IMPGMR_U   (3          )  010          HOW IMPORTANT IN GRADING? SPELL/PUNC/GRAMMAR: MOSTLY IMPORT
004   IMPGMR_?   (M,DNA      )  001          HOW IMPORT IN GRADING? SPELL/PUNC/GRAM: MISSING OR NOT APPLY

CONDITIONING ID:           TWRI0022
DESCRIPTION:               HOW IMPORTANT IN GRADING? ORGANIZATION
GRADES/ASSESSMENTS:        N08
GROUP LABEL:               IMP__ORG          LENGTH OF CONTRAST FIELD        :  3
NAEP ID:                   T049102           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:   4
001   IMPORG_V   (1          )  000          HOW IMPORTANT IN GRADING? ORGANIZATION: VERY IMPORTANT
002   IMPORG_M   (2          )  100          HOW IMPORTANT IN GRADING? ORGANIZATION: MODERATELY IMPORTANT
003   IMPORG_U   (3          )  010          HOW IMPORTANT IN GRADING? ORGANIZATION: MOSTLY IMPORTANT
004   IMPORG_?   (M,DNA      )  001          HOW IMPORTANT IN GRADING? ORGANIZATION: MISSING OR NOT APPLY

CONDITIONING ID:           TWRI0023
DESCRIPTION:               HOW IMPORTANT IN GRADING? CREATIVITY/QUALITY
GRADES/ASSESSMENTS:        N08
GROUP LABEL:               IMP_CREA          LENGTH OF CONTRAST FIELD        :  3
NAEP ID:                   T049103           DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:   4
001   IMPCRE_V   (1          )  000          HOW IMPORTANT IN GRADING? CREATIVITY/QUALITY: VERY IMPORTANT
002   IMPCRE_M   (2          )  100          HOW IMPORTANT IN GRADING? CREATIVITY/QUALITY: MODER IMPORT
```

Table F-4 (continued)
Conditioning Variables for Writing Main Samples

```
003   IMPCRE_U   (3          )  010         HOW IMPORTANT IN GRADING? CREATIVITY/QUALITY: MOSTLY IMPORT
004   IMPCRE_?   (M,DNA      )  001         HOW IMPORT IN GRADING? CREATIVITY/QUAL: MISSING OR NOT APPLY

CONDITIONING ID:              TWRI0024
DESCRIPTION:                  HOW IMPORTANT IN GRADING? LENGTH
GRADES/ASSESSMENTS:           N08
GROUP LABEL:                  IMP_LENG          LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                      T049104           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:             CLASS             NUMBER OF SPECIFICATION RECORDS:  4
001   IMPLEN_V   (1          )  000         HOW IMPORTANT IN GRADING? LENGTH: VERY IMPORTANT
002   IMPLEN_M   (2          )  100         HOW IMPORTANT IN GRADING? LENGTH: MODERATELY IMPORTANT
003   IMPLEN_U   (3          )  010         HOW IMPORTANT IN GRADING? LENGTH: MOSTLY IMPORTANT
004   IMPLEN_?   (M,DNA      )  001         HOW IMPORTANT IN GRADING? LENGTH: MISSING OR DOES NOT APPLY

CONDITIONING ID:              TWRI0025
DESCRIPTION:                  HOW IMPORTANT IN GRADING? ACCOMPLISH PURPOSE
GRADES/ASSESSMENTS:           N08
GROUP LABEL:                  IMP_PURP          LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                      T049105           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:             CLASS             NUMBER OF SPECIFICATION RECORDS:  4
001   IMPPUR_V   (1          )  000         HOW IMPORTANT IN GRADING? ACCOMPLISH PURPOSE: VERY IMPORTANT
002   IMPPUR_M   (2          )  100         HOW IMPORTANT IN GRADING? ACCOMP PURPOSE: MODERATE IMPORT
003   IMPPUR_U   (3          )  010         HOW IMPORTANT IN GRADING? ACCOMPLISH PURPOSE: MOSTLY IMPORT
004   IMPPUR_?   (M,DNA      )  001         HOW IMPORT IN GRADING? ACCOMP PURPOSE: MISSING OR NOT APPLY

CONDITIONING ID:              TWRI0026
DESCRIPTION:                  HOW OFTEN DO YOU ASK STUDENTS TO PLAN WRITING
GRADES/ASSESSMENTS:           N08
GROUP LABEL:                  ASK_PLAN          LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                      T049201           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:             CLASS             NUMBER OF SPECIFICATION RECORDS:  4
001   ASKPLN_A   (1          )  000         HOW OFTEN ASK STUDENTS TO PLAN WRITING: ALWAYS
002   ASKPLN_S   (2          )  100         HOW OFTEN ASK STUDENTS TO PLAN WRITING: SOMETIMES
003   ASKPLN_N   (3          )  010         HOW OFTEN ASK STUDENTS TO PLAN WRITING: NEVER
004   ASKPLN_?   (M,DNA      )  001         HOW OFTEN ASK STUDENTS TO PLAN WRITING: MISSING OR NOT APPLY

CONDITIONING ID:              TWRI0027
DESCRIPTION:                  HOW OFTEN DO YOU ASK STUDENTS TO OUTLINE GRADES/ASSESSMENTS:        N08
GROUP LABEL:                  ASK_OUT           LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                      T049202           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:             CLASS             NUMBER OF SPECIFICATION RECORDS:  4
001   ASKOUT_A   (1          )  000         HOW OFTEN ASK STUDENTS TO OUTLINE: ALWAYS
002   ASKOUT_S   (2          )  100         HOW OFTEN ASK STUDENTS TO OUTLINE: SOMETIMES
003   ASKOUT_N   (3          )  010         HOW OFTEN ASK STUDENTS TO OUTLINE: NEVER
004   ASKOUT_?   (M,DNA      )  001         HOW OFTEN ASK STUDENTS TO OUTLINE: MISSING OR DOES NOT APPLY

CONDITIONING ID:              TWRI0028
DESCRIPTION:                  HOW OFTEN DO YOU ASK STUDENTS TO DEFINE PURPOSE OR AUDIENCE
GRADES/ASSESSMENTS:           N08
GROUP LABEL:                  ASK_AUD           LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                      T049203           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:             CLASS             NUMBER OF SPECIFICATION RECORDS:  4
001   ASKAUD_A   (1          )  000         HOW OFTEN ASK STUDENTS TO DEFINE PURPOSE/AUDIENCE: ALWAYS
002   ASKAUD_S   (2          )  100         HOW OFTEN ASK STUDENTS TO DEFINE PURPOSE/AUDIENCE: SOMETIMES
003   ASKAUD_N   (3          )  010         HOW OFTEN ASK STUDENTS TO DEFINE PURPOSE/AUDIENCE: NEVER
004   ASKAUD_?   (M,DNA      )  001         HOW OFTEN ASK STUDENTS TO DEF PURP/AUD: MISSING OR NOT APPLY

CONDITIONING ID:              TWRI0029
DESCRIPTION:                  HOW OFTEN DO YOU ASK STUDENTS TO DISCUSS WRITING WITH STUDENTS
GRADES/ASSESSMENTS:           N08
GROUP LABEL:                  ASK_DSCS         .LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                      T049204           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:             CLASS             NUMBER OF SPECIFICATION RECORDS:  4
001   ASKDSC_A   (1          )  000         HOW OFTEN ASK TO DISCUSS WRITING WITH STUDENTS: ALWAYS
002   ASKDSC_S   (2          )  100         HOW OFTEN ASK TO DISCUSS WRITING WITH STUDENTS: SOMETIMES
```

Table F-4 (continued)
Conditioning Variables for Writing Main Samples

```
003   ASKDSC_N   (3          )  010         HOW OFTEN ASK TO DISCUSS WRITING WITH STUDENTS: NEVER
004   ASKDSC_?   (M,DNA      )  001         HOW OFTEN ASK TO DISCUSS WRITING WITH STUDENTS: MISSING OR

CONDITIONING ID:                TWRI0030
DESCRIPTION:                    HOW OFTEN DO YOU ASK STUDENTS TO DISCUSS WRITING WITH FAMILY
GRADES/ASSESSMENTS:             N08
GROUP LABEL:                    ASK_DSCF            LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                        T049205             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:               CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   ASKDSF_A   (1          )  000         HOW OFTEN ASK TO DISCUSS WRITING WITH FAMILY: ALWAYS
002   ASKDSF_S   (2          )  100         HOW OFTEN ASK TO DISCUSS WRITING WITH FAMILY: SOMETIMES
003   ASKDSF_N   (3          )  010         HOW OFTEN ASK TO DISCUSS WRITING WITH FAMILY: NEVER
004   ASKDSF_?   (M,DNA      )  001         HOW OFTEN ASK TO DISCUSS WRITING WITH FAMILY: MISSING OR DO

CONDITIONING ID:                TWRI0031
DESCRIPTION:                    HOW OFTEN DO YOU ASK STUDENTS TO WRITE MORE THAN 1 DRAFT
GRADES/ASSESSMENTS:             N08
GROUP LABEL:                    ASK_DRFT            LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                        T049206             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:               CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   ASKDFT_A   (1          )  000         HOW OFTEN ASK TO WRITE MORE THAN 1 DRAFT: ALWAYS
002   ASKDFT_S   (2          )  100         HOW OFTEN ASK TO WRITE MORE THAN 1 DRAFT: SOMETIMES
003   ASKDFT_N   (3          )  010         HOW OFTEN ASK TO WRITE MORE THAN 1 DRAFT: NEVER
004   ASKDFT_?   (M,DNA      )  001         HOW OFTEN ASK TO WRITE MORE THAN 1 DRAFT: MISSING OR DOES N

CONDITIONING ID:                TWRI0032
DESCRIPTION:                    HOW OFTEN DO YOU ASK STUDENTS TO COMMENT ON OTHERS' WRITINGS
GRADES/ASSESSMENTS:             N08
GROUP LABEL:                    ASK_COM             LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                        T049207             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:               CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   ASKCOM_A   (1          )  000         HOW OFTEN ASK TO COMMENT ON OTHERS' WRITINGS: ALWAYS
002   ASKCOM_S   (2          )  100         HOW OFTEN ASK TO COMMENT ON OTHERS' WRITINGS: SOMETIMES
003   ASKCOM_N   (3          )  010         HOW OFTEN ASK TO COMMENT ON OTHERS' WRITINGS: NEVER
004   ASKCOM_?   (M,DNA      )  001         HOW OFTEN ASK TO COMMENT ON OTHERS' WRITINGS: MISSING OR DO

CONDITIONING ID:                TWRI0033
DESCRIPTION:                    HOW OFTEN DO YOU ASK STUDENTS TO USE OTHER RESOURCES
GRADES/ASSESSMENTS:             N08
GROUP LABEL:                    ASK_RES             LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                        T049208             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:               CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   ASKRES_A   (1          )  000         HOW OFTEN ASK TO USE OTHER RESOURCES: ALWAYS
002   ASKRES_S   (2          )  100         HOW OFTEN ASK TO USE OTHER RESOURCES: SOMETIMES
003   ASKRES_N   (3          )  010         HOW OFTEN ASK TO USE OTHER RESOURCES: NEVER
004   ASKRES_?   (M,DNA      )  001         HOW OFTEN ASK TO USE OTHER RESOURCES: MISSING OR DOES NOT A

CONDITIONING ID:                TWRI0034
DESCRIPTION:                    HOW OFTEN DO YOU ASK STUDENTS TO TALK WITH YOU ABOUT WRITING
GRADES/ASSESSMENTS:             N08
GROUP LABEL:                    ASK_TLKT            LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                        T049209             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:               CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   ASKTLK_A   (1          )  000         HOW OFTEN ASK TO TALK WITH YOU ABOUT WRITING: ALWAYS
002   ASKTLK_S   (2          )  100         HOW OFTEN ASK TO TALK WITH YOU ABOUT WRITING: SOMETIMES
003   ASKTLK_N   (3          )  010         HOW OFTEN ASK TO TALK WITH YOU ABOUT WRITING: NEVER
004   ASKTLK_?   (M,DNA      )  001         HOW OFTEN ASK TO TALK WITH YOU ABOUT WRITING: MISSING OR DO

CONDITIONING ID:                TWRI0035
DESCRIPTION:                    HOW OFTEN DO YOU ASK STUDENTS TO CHOOSE WRITING TOPIC
GRADES/ASSESSMENTS:             N08
GROUP LABEL:                    ASK_CHOT            LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                        T049210             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:               CLASS               NUMBER OF SPECIFICATION RECORDS:  4
001   ASKCHO_A   (1          )  000         HOW OFTEN ASK TO CHOOSE WRITING TOPIC: ALWAYS
```

Table F-4 (continued)
Conditioning Variables for Writing Main Samples

```
002   ASKCHO_S   (2          )  100         HOW OFTEN ASK TO CHOOSE WRITING TOPIC: SOMETIMES
003   ASKCHO_N   (3          )  010         HOW OFTEN ASK TO CHOOSE WRITING TOPIC: NEVER
004   ASKCHO_?   (M,DNA      )  001         HOW OFTEN ASK TO CHOOSE WRITING TOPIC: MISSING OR DOES NOT

CONDITIONING ID:            TWRI0036
DESCRIPTION:                HOW OFTEN DO YOU ASK STUDENTS TO CONTRIBUTE TO A COLLECTION
GRADES/ASSESSMENTS:         N08
GROUP LABEL:                ASK_CNTB            LENGTH OF CONTRAST FIELD       :   3
NAEP ID:                    T049211             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   4
001   ASKCTB_A   (1          )  000         HOW OFTEN ASK TO CONTRIBUTE TO A COLLECTION: ALWAYS
002   ASKCTB_S   (2          )  100         HOW OFTEN ASK TO CONTRIBUTE TO A COLLECTION: SOMETIMES
003   ASKCTB_N   (3          )  010         HOW OFTEN ASK TO CONTRIBUTE TO A COLLECTION: NEVER
004   ASKCTB_?   (M,DNA      )  001         HOW OFTEN ASK TO CONTRIBUTE TO A COLLECTION: MISSING OR DOE

CONDITIONING ID:            TWRI0037
DESCRIPTION:                HOW OFTEN DO YOU ASK STUDENTS TO DO SPELLING OR GRAMMAR
GRADES/ASSESSMENTS:         N08
GROUP LABEL:                DO_SPELL            LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                    T049301             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   5
001   DOSP_DAY   (1          )  0000        HOW OFTEN ASK TO DO SPELLING OR GRAMMAR: ALMOST EVERY DAY
002   DOSP__WK   (2          )  1000        HOW OFTEN ASK TO DO SPELLING OR GRAMMAR: ONCE OR TWICE A WE
003   DOSP_MON   (3          )  0100        HOW OFTEN ASK TO DO SPELLING OR GRAMMAR: ONCE OR TWICE A MO
004   DOSP_NEV   (4          )  0010        HOW OFTEN ASK TO DO SPELLING OR GRAMMAR: NEVER OR HARDLY EV
005   DOSPEL_?   (M,DNA      )  0001        HOW OFTEN ASK TO DO SPELLING OR GRAMMAR: MISSING OR DOES N

CONDITIONING ID:            TWRI0038
DESCRIPTION:                HOW OFTEN DO YOU ASK STUDENTS TO WORK IN SMALL GROUPS
GRADES/ASSESSMENTS:         N08
GROUP LABEL:                DO_GRPS             LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                    T049302             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   5
001   DOGP_DAY   (1          )  0000        HOW OFTEN ASK TO WORK IN SMALL GROUPS: ALMOST EVERY DAY
002   DOGP__WK   (2          )  1000        HOW OFTEN ASK TO WORK IN SMALL GROUPS: ONCE OR TWICE A WEEK
003   DOGP_MON   (3          )  0100        HOW OFTEN ASK TO WORK IN SMALL GROUPS: ONCE OR TWICE A MONT
004   DOGP_NEV   (4          )  0010        HOW OFTEN ASK TO WORK IN SMALL GROUPS: NEVER OR HARDLY EVER
005   DOGRPS_?   (M,DNA      )  0001        HOW OFTEN ASK TO WORK IN SMALL GROUPS: MISSING OR DOES NOT

CONDITIONING ID:            TWRI0039
DESCRIPTION:                HOW OFTEN DO YOU ASK STUDENTS TO WRITE IN JOURNALS
GRADES/ASSESSMENTS:         N08
GROUP LABEL:                DO_JOUR             LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                    T049303             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   5
001   DOJR_DAY   (1          )  0000        HOW OFTEN ASK TO WRITE IN JOURNALS: ALMOST EVERY DAY
002   DOJR__WK   (2          )  1000        HOW OFTEN ASK TO WRITE IN JOURNALS: ONCE OR TWICE A WEEK
003   DOJR_MON   (3          )  0100        HOW OFTEN ASK TO WRITE IN JOURNALS: ONCE OR TWICE A MONTH
004   DOJR_NEV   (4          )  0010        HOW OFTEN ASK TO WRITE IN JOURNALS: NEVER OR HARDLY EVER
005   DOJOUR_?   (M,DNA      )  0001        HOW OFTEN ASK TO WRITE IN JOURNALS: MISSING OR DOES NOT AP

CONDITIONING ID:            TWRI0040
DESCRIPTION:                HOW OFTEN DO YOU ASSESS PROGRESS WITH A MULTIPLE CHOICE TEST
GRADES/ASSESSMENTS:         N08
GROUP LABEL:                ASSES_MC            LENGTH OF CONTRAST FIELD       :   4
NAEP ID:                    T049501             DEGREES OF FREEDOM PER CONTRAST:   1
TYPE OF CONTRAST:           CLASS               NUMBER OF SPECIFICATION RECORDS:   5
001   ASMC_DAY   (1          )  0000        HOW OFTEN ASSESS WITH A MULTIPLE CHOICE TEST: ALMOST EVERY
002   ASMC__WK   (2          )  1000        HOW OFTEN ASSESS WITH A MULTIPLE CHOICE TEST: ONCE OR TWICE
003   ASMC_MON   (3          )  0100        HOW OFTEN ASSESS WITH A MULTIPLE CHOICE TEST: ONCE OR TWICE
004   ASMC_NEV   (4          )  0010        HOW OFTEN ASSESS WITH A MULTIPLE CHOICE TEST: NEVER OR HARD
005   ASSEMC_?   (M,DNA      )  0001        HOW OFTEN ASSESS WITH A MULTIPLE CHOICE TEST: MISSING OR D

CONDITIONING ID:            TWRI0041
DESCRIPTION:                HOW OFTEN DO YOU ASSESS PROGRESS WITH A SHORT ANSWER TEST
```

Table F-4 (continued)
Conditioning Variables for Writing Main Samples

```
GRADES/ASSESSMENTS:        N08
GROUP LABEL:               ASSES_SH          LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                   T049502           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:  5
001   ASSH_DAY   (1          )  0000         HOW OFTEN ASSESS WITH A SHORT ANSWER TEST: ALMOST EVERY DAY
002   ASSH__WK   (2          )  1000         HOW OFTEN ASSESS WITH A SHORT ANSWER TEST: ONCE OR TWICE A
003   ASSH_MON   (3          )  0100         HOW OFTEN ASSESS WITH A SHORT ANSWER TEST: ONCE OR TWICE A
004   ASSH_NEV   (4          )  0010         HOW OFTEN ASSESS WITH A SHORT ANSWER TEST: NEVER OR HARDLY
005   ASSESH_?   (M,DNA      )  0001         HOW OFTEN ASSESS WITH A SHORT ANSWER TEST: MISSING OR DOES

CONDITIONING ID:           TWRI0042
DESCRIPTION:               HOW OFTEN DO YOU ASSESS PROGRESS WITH LONG ESSAYS
GRADES/ASSESSMENTS:        N08
GROUP LABEL:               ASSES_ES          LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                   T049503           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:  5
001   ASES_DAY   (1          )  0000         HOW OFTEN ASSESS WITH LONG ESSAYS: ALMOST EVERY DAY
002   ASES__WK   (2          )  1000         HOW OFTEN ASSESS WITH LONG ESSAYS: ONCE OR TWICE A WEEK
003   ASES_MON   (3          )  0100         HOW OFTEN ASSESS WITH LONG ESSAYS: ONCE OR TWICE A MONTH
004   ASES_NEV   (4          )  0010         HOW OFTEN ASSESS WITH LONG ESSAYS: NEVER OR HARDLY EVER
005   ASSEES_?   (M,DNA      )  0001         HOW OFTEN ASSESS WITH LONG ESSAYS: MISSING OR DOES NOT APP

CONDITIONING ID:           TWRI0043
DESCRIPTION:               HOW OFTEN DO YOU ASSESS PROGRESS WITH PROJECTS OR PORTFOLIOS
GRADES/ASSESSMENTS:        N08
GROUP LABEL:               ASSES_PF          LENGTH OF CONTRAST FIELD       :  4
NAEP ID:                   T049504           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:  5
001   ASPF_DAY   (1          )  0000         HOW OFTEN ASSESS WITH PROJECTS/PORTFOLIOS: ALMOST EVERY DAY
002   ASPF__WK   (2          )  1000         HOW OFTEN ASSESS WITH PROJECTS/PORTFOLIOS: ONCE OR TWICE/WK
003   ASPF_MON   (3          )  0100         HOW OFTEN ASSESS WITH PROJECTS/PORTFOLIOS: ONCE OR TWICE/MON
004   ASPF_NEV   (4          )  0010         HOW OFTEN ASSESS WITH PROJECTS/PORTFOLIOS: NEVER OR HARDLY
005   ASSEPF_?   (M,DNA      )  0001         HOW OFTEN ASSESS WITH PROJ/PORTFOLIOS: MISSING OR NOT APPLY

CONDITIONING ID:           TWRI0044
DESCRIPTION:               HOW OFTEN DO YOU ASK STUDENTS TO DO EXERCISES ON COMPUTERS
GRADES/ASSESSMENTS:        N08
GROUP LABEL:               TEX_COMP          LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                   T049601           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:  4
001   EXCP__WK   (1          )  000          HOW OFTEN DO EXERCISES ON COMPUTERS: ONCE OR TWICE A
002   EXCP_MON   (2          )  100          HOW OFTEN DO EXERCISES ON COMPUTERS: ONCE OR TWICE A
003   EXCP_NEV   (3          )  010          HOW OFTEN DO EXERCISES ON COMPUTERS: NEVER OR HARDLY
004   EXCOMP_?   (M,DNA      )  001          HOW OFTEN DO EXERCISES ON COMPUTERS: MISSING OR DOE

CONDITIONING ID:           TWRI0045
DESCRIPTION:               HOW OFTEN DO YOU ASK STUDENTS TO DO WRITING ON COMPUTERS
GRADES/ASSESSMENTS:        N08
GROUP LABEL:               TWT_COMP          LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                   T049602           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:  4
001   WTCP__WK   (1          )  000          HOW OFTEN ASK STUDENTS TO WRITING COMPUTERS: ONCE, TWICE W
002   WTCP_MON   (2          )  100          HOW OFTEN ASK STUDENTS TO WRITING COMPUTERS: ONCE,TWICE M
003   WTCP_NEV   (3          )  010          HOW OFTEN ASK STUDENTS TO WRITING ON COMPUTERS: NEVER
004   WTCOMP_?   (M,DNA      )  001          HOW OFTEN ASK STUDENTS TO WRITING ON COMPUTERS: MISSING

CONDITIONING ID:           TWRI0046
DESCRIPTION:               ARE COMPUTERS AVAILABLE FOR YOUR WRITING CLASS
GRADES/ASSESSMENTS:        N08
GROUP LABEL:               TAV_COMP          LENGTH OF CONTRAST FIELD       :  3
NAEP ID:                   T049701           DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:          CLASS             NUMBER OF SPECIFICATION RECORDS:  4
001   AVCP_NOT   (1          )  000          ARE COMPUTERS AVAILABLE FOR CLASS: NOT AVAILABLE
002   AVCP_YDF   (2          )  100          ARE COMPUTERS AVAILABLE FOR CLASS: AVAILABLE IN DIFFERENT
003   AVCP_YCL   (3          )  010          ARE COMPUTERS AVAILABLE FOR CLASS: AVAILABLE IN CLASSROOM
```

BEST COPY AVAILABLE

Table F-4 (continued)
Conditioning Variables for Writing Main Samples

```
004   AVCOMP_?   (M,DNA       )  001         ARE COMPUTERS AVAILABLE FOR CLASS: MISSING OR NOT APPLY

CONDITIONING ID:             TWRI0047
DESCRIPTION:                 TEACHER: IS STUDENTS' WRITING KEPT IN A PORTFOLIO?
GRADES/ASSESSMENTS:          N08
GROUP LABEL:                 TPORTFOL            LENGTH OF CONTRAST FIELD       :  2
NAEP ID:                     T049401             DEGREES OF FREEDOM PER CONTRAST:  1
TYPE OF CONTRAST:            CLASS               NUMBER OF SPECIFICATION RECORDS:  3
001   TPORT__Y   (1           )  00         TEACHER: IS STUDENTS' WRITING KEPT IN A PORTFOLIO: YES
002   TPORT__N   (2           )  10         TEACHER: IS STUDENTS' WRITING KEPT IN A PORTFOLIO: NO
003   TPORT__?   (M,DNA       )  01         TEACHER: IS STUDENTS' WRITING KEPT IN A PORTFOLIO: MISSING
```

Table F-5
Conditioning Variables for the Reading Trend Samples

Conditioning Variable	Age Classes	Variable Name(s)	Variable Coding	Contrast Coding*
Overall	All		---	1
Gender	All	DSEX	Male	0
			Female	1
Size and Type of Community	All	STOC	Low Metro	00
			High Metro	10
			All others and Missing	01
Region	All	REGION	Northeast	000
			Southeast	100
			Central	010
			West	001
Parents' Education	All	PARED	Less than high school	0000
			High school graduate	1000
			Post-high school	0100
			College graduate	0010
			Missing and I Don't Know	0001
Items in the Home	All	B000901	None of the six items	00
		B000902	One of the six items	10
		B000903	Two of the six items	20
		B000904	Three of the six items	30
		B000905	Four of the six items	40
		B000906	Five of the six items	50
			Six of the six items	60
			Missing	01
Television Watching	All	B001801	None	00
			One hour or less	10
			Two hours	20
			Three hours	30
			Four hours	40
			Five hours	50
			Six or more hours	60
			Missing	01
Homework	All	B001701	Don't have any	00
			Don't do any	00
			Less than 1 hour	10
			1-2 hours	20
			More than 2 hours	30
			Missing	01

* Multicolumn entries without overbars indicate multiple contrasts. Barred columns are treated as one contrast.

Table F-5 (continued)
Conditioning Variables for the Reading Trend Samples

Conditioning Variable	Age Classes	Variable Name(s)	Variable Coding	Contrast Coding*
Language Spoken at Home	All	B000401	English	00
			Spanish	10
			Other	10
			Missing	01
Pages Read	All	B001101	More than 20	10
			16-20	10
			11-15	10
			6-10	10
			5 or fewer	00
			Missing	01
Percent in School Lunch Program	All	PCLUNCH	0 percent	$\overline{000}$ 0
			1 percent	001 0
			2 percent	002 0
			.	.
			.	.
			99 percent	099 0
			100 percent	100 0
			Missing	000 1
Percent White	All	PCTWHT	0-49 Minority	100
			50-79 Integrated	010
			80-100 Predominantly White	001
			Missing	000
Courses Taken	9, 13	B001001	None of the seven	$\overline{00.0}$ 0
		B001002	One of the seven	01.0 0
		B002003	Two of the seven	02.0 0
		B002004	Three of the seven	03.0 0
		B002005	Four of the seven	04.0 0
		B002006	Five of the seven	05.0 0
		B002007	Six of the seven	06.0 0
			Seven of the seven	07.0 0
			Missing	00.0 1
Derived Race/Ethnicity	All	DRACE	White	000
			Black	100
			Hispanic	010
			Asian American	001
			American Indian	000
			Unclassified	000
			Missing	000

* Multicolumn entries without overbars indicate multiple contrasts. Barred columns are treated as one contrast.

Table F-5 (continued)
Conditioning Variables for the Reading Trend Samples

Conditioning Variable	**Age Classes**	**Variable Name(s)**	**Variable Coding**	**Contrast Coding***
Age by Grade	All	MODGRD MODAGE	1 < modal age, modal grade	0000
			2 Modal age, < modal grade	1000
			3 Modal age, modal grade/missing	0100
			4 Modal age, > modal grade	0010
			5 > modal age, modal grade	0001

* Multicolumn entries without overbars indicate multiple contrasts. Barred columns are treated as one contrast.

Table F-6
Conditioning Variables for the Mathematics Trend Samples

Conditioning Variable	Age Classes	Variable Name(s)	Variable Coding	Contrast Coding*
Overall	All		---	1
Gender	All	DSEX	Male	0
			Female	1
Observed Race/Ethnicity	All	RACE	White	000
			Black	100
			Hispanic	010
			Asian American	001
			American Indian	000
			Other	000
			Missing	000
Size and Type of Community	All	STOC	1, 4-7 all except 2 and 3	01
			2 Low Metro	00
			3 High Metro	10
Region	All	REGION	Northeast	000
			Southeast	100
			Central	010
			West	001
Parents' Education	All	PARED	Less than high school	0000
			High school graduate	1000
			Post-high school	0100
			College graduate	0010
			Missing and I Don't Know	0001
Modal Grade	All	MODGRD	< modal grade	10
			= modal grade, missing	00
			> modal grade	01
Items in the Home (of newspaper, > 25 books, encyclopedia, magazines)	All	HOMEEN2	0 to 2 items	00
			3 items	10
			4 items	01

---

* Multicolumn entries without overbars indicate multiple contrasts. Barred columns are treated as one contrast.

Table F-6 (continued)
Conditioning Variables for the Mathematics Trend Samples

Conditioning Variable	**Age Classes**	**Variable Name(s)**	**Variable Coding**	**Contrast Coding***
Observed Race/Ethnicity by Gender ("White" includes American Indian and Other)	All	RACE DSEX	White, male	000
			Black, male	000
			Hispanic, male	000
			Asian American, male	000
			White, female	000
			Black, female	100
			Hispanic, female	010
			Asian American, female	001
Observed Race/Ethnicity by Parents' Education ("White" includes American Indian and Other)--coded differently for each age class	9	RACE PARED	White, < HS	0000 0000 0000
			White, HS graduate	0000 0000 0000
			White, post-HS	0000 0000 0000
			White, college grad.	0000 0000 0000
			White, missing	0000 0000 0000
			Black, < HS	0000 0000 0000
			Black, HS grad & post-HS	1000 0000 0000
			Black, college grad.	0010 0000 0000
			Black, missing	0001 0000 0000
			Hispanic, < HS	0000 0000 0000
			Hispanic, HS grad & post-HS	0000 1000 0000
			Hispanic, coll. grad.	0000 0010 0000
			Hispanic, missing	0000 0001 0000
			Asian Amer., < HS	0000 0000 0000
			Asian Amer., HS grad & post-HS	0000 0000 1000
			Asian Amer., coll. grad.	0000 0000 0010
			Asian Amer., missing	0000 0000 0001

---

* Multicolumn entries without overbars indicate multiple contrasts. Barred columns are treated as one contrast.

Table F-6 (continued)
Conditioning Variables for the Mathematics Trend Samples

Conditioning Variable	Age Classes	Variable Name(s)	Variable Coding	Contrast Coding*
Observed Race/Ethnicity by Parents' Education ("White" includes American Indian and Other)--coded differently for each age class	13	RACE PARED	White, < HS	0000 0000 0000
			White, HS graduate	0000 0000 0000
			White, post-HS	0000 0000 0000
			White, college grad.	0000 0000 0000
			White, missing	0000 0000 0000
			Black, < HS	0000 0000 0000
			Black, HS graduate	1000 0000 0000
			Black, post-HS	0100 0000 0000
			Black, college grad.	0010 0000 0000
			Black, missing	0001 0000 0000
			Hispanic, < HS	0000 0000 0000
			Hispanic, HS grad.	0000 1000 0000
			Hispanic, post-HS	0000 0100 0000
			Hispanic, coll. grad.	0000 0010 0000
			Hispanic, missing	0000 0001 0000
			Asian Amer., < HS	0000 0000 0000
			Asian Amer., HS grad.	0000 0000 1000
			Asian Amer., post-HS	0000 0000 0100
			Asian Amer., coll. grad.	0000 0000 0010
			Asian Amer., missing	0000 0000 0001
Observed Race/Ethnicity by Parents' Education ("White" includes American Indian and Other)--coded differently for each age class	17	RACE PARED	White, < HS	0000 0000 0000
			White, HS graduate	0000 0000 0000
			White, post-HS	0000 0000 0000
			White, college grad.	0000 0000 0000
			White, missing	0000 0000 0000
			Black, < HS	0000 0000 0000
			Black, HS graduate	1000 0000 0000
			Black, post-HS	0100 0000 0000
			Black, college grad.	0010 0000 0000
			Black, missing	0001 0000 0000
			Hispanic, < HS	0000 0000 0000
			Hispanic, HS grad.	0000 1000 0000
			Hispanic, post-HS	0000 0100 0000
			Hispanic, coll. grad.	0000 0010 0000
			Hispanic, missing	0000 0001 0000
			Asian Amer., < HS	0000 0000 0000
			Asian Amer., HS grad.	0000 0000 1000
			Asian Amer., post-HS, coll. grad.	0000 0000 0100
			Asian Amer., missing	0000 0000 0001

---

* Multicolumn entries without overbars indicate multiple contrasts. Barred columns are treated as one contrast.

Table F-6 (continued)
Conditioning Variables for the Mathematics Trend Samples

Conditioning Variable	Age Classes	Variable Name(s)	Variable Coding	Contrast Coding*
School Type	All	SCHTYPE	Public	0
			Private	1
			Catholic	1
			Bureau of Indian Affairs	1
			Department of Defense	1
Homework	13, 17	B003901	None assigned	100
			Didn't do	010
			½ hour or less	012
			1 hour	013
			2 hours	014
			More than 2 hours	000
			Missing	000
Language in the Home	All	LANGHOM	Never	00
			Sometimes	10
			Always	01
Observed Race/Ethnicity by Language in the Home--coded differently for age class 9	9	RACE LANGHOM	White, often	00 00 00
			White, sometimes	00 00 00
			White, never	00 00 00
			Black, often & sometimes	10 00 00
			Black, never	00 00 00
			Hispanic, often & sometimes	00 10 00
			Hispanic, never	00 00 00
			Asian Amer., often & sometimes	00 00 10
			Asian Amer., never	00 00 00
Observed Race/Ethnicity by Language in the Home	13, 17	RACE LANGHOM	White, often	00 00 00
			White, sometimes	00 00 00
			White, never	00 00 00
			Black, often	10 00 00
			Black, sometimes	01 00 00
			Black, never	00 00 00
			Hispanic, often	00 10 00
			Hispanic, sometimes	00 01 00
			Hispanic, never	00 00 00
			Asian Amer., often	00 00 10
			Asian Amer., sometimes	00 00 01
			Asian Amer., never	00 00 00

* Multicolumn entries without overbars indicate multiple contrasts. Barred columns are treated as one contrast.

Table F-6 (continued)
Conditioning Variables for the Mathematics Trend Samples

Conditioning Variable	Age Classes	Variable Name(s)	Variable Coding	Contrast Coding*
Highest Level Math Taken	17	NMATH	Pre-algebra	10000
			Algebra	01000
			Geometry	00100
			Algebra 2	00010
			Calculus	00001
			Something else	00000
High School Program	17	B005001	General	00
			College preparatory	10
			Vocational, technical	01
			Missing	00
Derived Race/Ethnicity	All	DRACE	White	000
			Black	100
			Hispanic	010
			Asian American	001
			Other	000
			Missing	000

---

* Multicolumn entries without overbars indicate multiple contrasts. Barred columns are treated as one contrast.

Table F-7
Conditioning Variables for the Writing Trend Samples

Conditioning Variable	Age Classes	Variable Name(s)	Variable Coding	Contrast Coding*
Overall	All		---	1
Gender	All	DSEX	Male	0
			Female	1
Race/Ethnicity	All	DRACE	White	10
			Black	00
			Hispanic	01
			Asian American	10
			American Indian	10
			Unclassified	10
			Missing	10
Size and Type of Community	All	STOC	2 Low Metro	10
			3 High Metro	01
			1, 4-7 All others and Missing	00
Region	All	REGION	Northeast	000
			Southeast	100
			Central	010
			West	001
Parents' Education	All	PARED	Less than high school	0000
			High school graduate	1000
			Post-high school	0100
			College graduate	0010
			Missing and I Don't Know	0001
Modal Grade	All	MODGRD	< modal grade	10
			= modal grade, missing	00
			> modal grade	01
Articles in Home	All	B000901 to B000905	0 to 3 articles	00
			4 articles	10
			5 articles	01
			Missing	00
School Type	All	SCHTYPE	Public	0
			Private	1
			Catholic	1
			Bureau of Indian Affairs	1
			Department of Defense	1
			Missing	1

---

* Multicolumn entries without overbars indicate multiple contrasts. Barred columns are treated as one contrast.

873

Table F-7 (continued)
Conditioning Variables for the Writing Trend Samples

Conditioning Variable	Age Classes	Variable Name(s)	Variable Coding	Contrast Coding*
Homework	All	B001701	None assigned	00000
			Didn't do	10000
			< 1 hour	01000
			1 to 2 hours	00100
			> 2 hours	00010
			Missing	00001
TV Watching	All	B001801	None	0 $\overline{00}$
			1 hour or less	1 01
			2 hours	2 04
			3 hours	3 09
			4 hours	4 16
			5 hours	5 25
			6 hours or more	6 36
			Missing	3 09
Mother Works Outside Home	All	B000801	Not work	0
			Works	1
			Missing	0
Language Minority	All	B000301	English	00
			Spanish	10
			Other	10
			Missing	01
Percent in Lunch Program	All	C003201	0 percent	0000000
			1 - 5%	1000000
			6 - 10%	0100000
			11 - 25%	0010000
			26 - 50%	0001000
			51 - 75%	0000100
			76 - 90%	0000010
			> 90%	0000001
			Missing	0000000
Percent White in School	All	PCTWHT	0 - 49.9% minority	00
			50 - 79.9 integrated	10
			80 - 100% predominantly white	01
			Missing	01

---

* Multicolumn entries without overbars indicate multiple contrasts. Barred columns are treated as one contrast.

Table F-7 (continued)
Conditioning Variables for the Writing Trend Samples

Conditioning Variable	**Age Classes**	**Variable Name(s)**	**Variable Coding**	**Contrast Coding***
Grades in School	All	B001901	Mostly A	$\overline{4.0}$
			Mostly B	3.0
			Mostly C	2.0
			Mostly D	1.0
			Less than D	0.0
			Between A and B	3.5
			Between B and C	2.5
			Between C and D	1.5
			Between D and F	0.5
			Missing	2.0
Pages a Day Read for School and Homework	All	B001101	More than 20	100
			16 to 20	100
			11 - 15	100
			6 - 10	010
			5 or fewer	001
			Missing	000
Number of Essays, Reports Written	All	B001201 to B001208	0 reports	0
			1 report	1
			2 reports	2
			3 reports	3
			4 reports	4
			5 reports	5
			6 reports	6
			7 reports	7
			Missing	0

---

* Multicolumn entries without overbars indicate multiple contrasts. Barred columns are treated as one contrast.

Table F-8
Conditioning Variables for the Science Trend Samples

Conditioning Variable	Age Classes	Variable Name(s)	Variable Coding	Contrast Coding*
Overall	All		---	1
Gender	All	DSEX	Male Female	0 1
Observed Race/Ethnicity	All	RACE	White Black Hispanic Asian American American Indian Other Missing	000 100 010 001 000 000 000
Size and Type of Community	All	STOC	2 Low Metro 3 High Metro 1, 4-7 All others and Missing	10 01 00
Region	All	REGION	Northeast Southeast Central West	000 100 010 001
Parents' Education	All	PARED	Less than high school High school graduate Post-high school College graduate Missing and I Don't Know	0000 1000 0100 0010 0001
Modal Grade	All	MODGRD	< modal grade = modal grade, missing > modal grade	10 00 01
Items in the Home (of newspaper, > 25 books, encyclopedia, magazines)	All	HOMEEN2	0 to 2 items 3 items 4 items	00 10 01

---

* Multicolumn entries without overbars indicate multiple contrasts. Barred columns are treated as one contrast.

Table F-8 (continued)
Conditioning Variables for the Writing Trend Samples

Conditioning Variable	Age Classes	Variable Name(s)	Variable Coding	Contrast Coding*
Observed Race/Ethnicity by Gender ("White" includes American Indian and Other)	All	RACE DSEX	White, male	000
			Black, male	000
			Hispanic, male	000
			Asian American, male	000
			White, female	000
			Black, female	100
			Hispanic, female	010
			Asian American, female	001
			Missing	000
Observed Race/Ethnicity by Parents' Education ("White" includes American Indian and Other)	All	RACE PARED	White, < HS	0000 0000 0000
			White, HS graduate	0000 0000 0000
			White, post-HS	0000 0000 0000
			White, college grad.	0000 0000 0000
			White, missing	0000 0000 0000
			Black, < HS	0000 0000 0000
			Black, HS graduate	1000 0000 0000
			Black, post-HS	0100 0000 0000
			Black, college grad.	0010 0000 0000
			Black, missing	0001 0000 0000
			Hispanic, < HS	0000 0000 0000
			Hispanic, HS grad.	0000 1000 0000
			Hispanic, post-HS	0000 0100 0000
			Hispanic, coll. grad.	0000 0010 0000
			Hispanic, missing	0000 0001 0000
			Asian Amer., < HS	0000 0000 0000
			Asian Amer., HS grad.	0000 0000 1000
			Asian Amer., post-HS	0000 0000 0100
			Asian Amer., coll. grad.	0000 0000 0010
			Asian Amer., missing	0000 0000 0001
School Type	All	SCHTYPE	Public	0
			Private	1
			Catholic	1
			Bureau of Indian Affairs	1
			Department of Defense	1

* Multicolumn entries without overbars indicate multiple contrasts. Barred columns are treated as one contrast.

Table F-8 (continued)
Conditioning Variables for the Writing Trend Samples

Conditioning Variable	Age Classes	Variable Name(s)	Variable Coding	Contrast Coding*
Homework	13, 17	B003901	None assigned	100
			Didn't do	010
			½ hour or less	012
			1 hour	013
			2 hours	014
			More than 2 hours	000
			Missing	000
Language in the Home	All	LANGHOM	Never	00
			Sometimes	10
			Always	01
Observed Race/Ethnicity by Language in the Home	All	RACE LANGHOM	White, often	00 00 00
			White, sometimes	00 00 00
			White, never	00 00 00
			Black, often	10 00 00
			Black, sometimes	01 00 00
			Black, never	00 00 00
			Hispanic, often	00 10 00
			Hispanic, sometimes	00 01 00
			Hispanic, never	00 00 00
			Asian Amer., often	00 00 10
			Asian Amer., sometimes	00 00 01
			Asian Amer., never	00 00 00
Number of Science Courses	17	NSCI	General science	1000
			Biology	0100
			Chemistry	0010
			Physics	0001
			Nothing/something else	0000
High School Program	17	B005001	General	00
			College preparatory	10
			Vocational, technical	01
			Missing	00
Derived Race/Ethnicity	All	DRACE	White	000
			Black	100
			Hispanic	010
			Asian American	001
			Other	000
			Missing	000

---

* Multicolumn entries without overbars indicate multiple contrasts. Barred columns are treated as one contrast.

Table F-9
Estimated Effects for the Reading Main Principal Components
Grade 4, Reading for Literary Experience

Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect
0	-0.705617	45	-0.016459	90	-0.015411	135	-0.013616
1	-0.025113	46	0.012614	91	-0.004754	136	-0.001128
2	0.076504	47	0.016080	92	-0.015445	137	0.014421
3	-0.077738	48	-0.015904	93	-0.036504	138	-0.013458
4	0.018061	49	-0.000288	94	-0.021886	139	-0.011382
5	-0.010294	50	0.014711	95	0.008272	140	-0.023132
6	-0.114765	51	-0.005741	96	0.021272	141	0.020887
7	-0.060756	52	-0.020746	97	0.000601	142	0.012245
8	-0.013344	53	0.028888	98	0.005395	143	-0.007273
9	0.035079	54	-0.001599	99	-0.006034	144	-0.012974
10	-0.014852	55	0.006798	100	0.005647	145	-0.001675
11	0.039914	56	-0.004445	101	-0.005610	146	0.014454
12	0.040879	57	0.006164	102	-0.001802	147	0.040144
13	0.018994	58	0.005824	103	0.007949	148	0.030496
14	-0.045345	59	0.015888	104	0.006449	149	0.039221
15	0.007904	60	0.025727	105	0.019338	150	-0.018171
16	0.013591	61	-0.008709	106	-0.008839	151	0.013961
17	0.047111	62	0.015564	107	0.009279	152	-0.042895
18	-0.035375	63	0.013412	108	-0.015719	153	0.006450
19	0.020367	64	-0.018491	109	0.002336	154	-0.005551
20	-0.031914	65	0.008297	110	0.005178	155	0.032876
21	-0.050661	66	-0.023090	111	0.012743	156	0.014024
22	-0.060122	67	-0.032325	112	0.009327	157	-0.005479
23	-0.051150	68	0.005882	113	-0.019055	158	0.008400
24	-0.004762	69	-0.012685	114	0.015738	159	0.008961
25	0.037146	70	0.024037	115	0.004744	160	0.015472
26	0.001559	71	0.007654	116	0.006055	161	0.037401
27	0.023155	72	0.000456	117	0.012041	162	0.010479
28	0.042164	73	-0.006173	118	0.024148	163	0.048701
29	0.003322	74	0.004698	119	-0.006888	164	0.008871
30	-0.017679	75	0.004725	120	-0.030310	165	0.008908
31	0.041658	76	-0.006705	121	0.028995	166	0.021206
32	0.014474	77	-0.016357	122	-0.022001	167	-0.045064
33	0.033495	78	0.000325	123	0.000084	168	-0.012414
34	0.003853	79	0.005428	124	-0.026314	169	-0.013818
35	-0.011247	80	0.008399	125	0.006161		
36	0.009206	81	-0.001986	126	-0.020109		
37	-0.025640	82	-0.003070	127	-0.003067		
38	-0.015525	83	0.030583	128	-0.011195		
39	0.013133	84	0.034671	129	-0.010392		
40	-0.008042	85	-0.028572	130	-0.012403		
41	0.000803	86	0.006775	131	-0.021954		
42	-0.018820	87	-0.020479	132	-0.013452		
43	0.001153	88	0.005230	133	0.040734		
44	0.021177	89	0.004815	134	0.008521		

Table F-10
Estimated Effects for the Reading Main Principal Components
Grade 4, Reading to Gain Information

Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect
0	-1.009626	45	0.008410	90	-0.011536	135	-0.038931
1	-0.026856	46	0.009975	91	0.017790	136	-0.006601
2	0.089323	47	0.035112	92	-0.022207	137	0.013028
3	-0.080348	48	0.001877	93	-0.014786	138	-0.008969
4	0.012626	49	0.003739	94	-0.012426	139	-0.012034
5	-0.013376	50	0.024906	95	0.013368	140	-0.007634
6	-0.128031	51	-0.016938	96	0.002082	141	0.013755
7	-0.060620	52	-0.007545	97	-0.000134	142	0.024555
8	-0.020901	53	0.027316	98	0.000573	143	-0.012885
9	0.047277	54	-0.020978	99	0.006093	144	-0.017382
10	-0.023509	55	0.012751	100	0.000279	145	0.010938
11	0.016853	56	-0.003343	101	0.025196	146	-0.001297
12	0.055343	57	0.004792	102	0.001381	147	0.038842
13	0.025018	58	0.002431	103	0.004707	148	0.008167
14	-0.030424	59	0.020453	104	0.016455	149	0.010723
15	0.006784	60	0.025177	105	0.013611	150	-0.004640
16	0.008980	61	-0.012343	106	0.015494	151	-0.001767
17	0.032846	62	0.011322	107	0.002056	152	-0.018279
18	-0.033181	63	0.002059	108	-0.006115	153	0.057308
19	0.007258	64	-0.024031	109	0.038234	154	0.011778
20	-0.015083	65	0.004821	110	-0.003984	155	0.007561
21	-0.050506	66	0.004050	111	0.017770	156	0.015853
22	-0.077737	67	-0.043551	112	-0.011725	157	0.022921
23	-0.028660	68	0.007465	113	-0.016242	158	-0.011741
24	0.004544	69	-0.029537	114	-0.004390	159	-0.005053
25	0.036077	70	0.024083	115	0.031054	160	0.001341
26	0.005726	71	0.001568	116	-0.001903	161	0.038701
27	0.012790	72	-0.011714	117	0.020536	162	0.009277
28	0.043870	73	0.002325	118	-0.008538	163	0.041597
29	0.007533	74	0.039133	119	-0.030488	164	0.014224
30	-0.025198	75	0.013402	120	-0.022023	165	-0.010529
31	0.051089	76	-0.029395	121	0.031401	166	0.020104
32	0.005916	77	-0.017124	122	-0.006042	167	-0.017858
33	0.043717	78	-0.010439	123	0.004731	168	0.009939
34	0.001644	79	0.027936	124	0.002631	169	0.026663
35	-0.008118	80	0.001700	125	0.000262		
36	0.002453	81	0.023021	126	-0.032915		
37	-0.038541	82	0.004170	127	-0.004435		
38	-0.018192	83	0.026969	128	-0.040694		
39	0.012717	84	0.047761	129	-0.005008		
40	-0.016106	85	0.000536	130	-0.015613		
41	0.000595	86	-0.011501	131	-0.009362		
42	-0.020190	87	0.000283	132	-0.029063		
43	-0.009524	88	-0.023734	133	0.062617		
44	0.035795	89	0.033230	134	0.021010		

Table F-11
Estimated Effects for the Reading Main Principal Components
Grade 8, Reading for Literary Experience

Principal Component	Estimated Effect
0	0.100628
1	-0.048463
2	-0.006984
3	-0.072519
4	0.118164
5	-0.094289
6	0.043000
7	-0.049819
8	-0.009443
9	0.030830
10	0.100041
11	0.016874
12	-0.048207
13	0.008563
14	-0.027268
15	-0.015914
16	0.009395
17	-0.056471
18	-0.001287
19	-0.035553
20	-0.012252
21	0.056245
22	-0.004576
23	0.035371
24	0.038260
25	0.007019
26	0.015618
27	-0.029436
28	-0.011009
29	-0.016812
30	-0.040235
31	0.010920
32	0.019636
33	-0.006162
34	-0.005805
35	-0.002104
36	0.008566
37	0.003321
38	-0.006198
39	-0.017034
40	-0.022971
41	0.008851
42	-0.012576
43	0.000546
44	0.023577
45	-0.036931
46	0.015147
47	-0.020213
48	0.010258
49	0.021490
50	0.027177
51	-0.006494
52	-0.003474
53	0.032553
54	0.005010
55	-0.001467
56	-0.008900
57	-0.020715
58	-0.019441
59	-0.006836
60	0.027462
61	0.038169
62	-0.017492
63	-0.002073
64	0.016668
65	-0.026479
66	-0.009581
67	-0.005310
68	-0.006311
69	-0.004176
70	0.002832
71	0.015025
72	-0.012080
73	-0.023637
74	-0.008809
75	0.010081
76	0.002765
77	-0.003864
78	-0.001232
79	-0.015442
80	0.022436
81	0.022603
82	0.035263
83	-0.020719
84	0.002047
85	0.002849
86	-0.004779
87	-0.031926
88	0.002289
89	-0.024442
90	-0.010806
91	-0.009921
92	0.008196
93	0.021797
94	-0.025636
95	0.000417
96	-0.012132
97	-0.023137
98	0.010419
99	-0.028970
100	-0.011079
101	0.068721
102	-0.034646
103	-0.051292
104	-0.011130
105	0.000748
106	-0.014099
107	-0.022875
108	0.004054
109	-0.018409
110	0.005164
111	-0.023162

Table F-12
Estimated Effects for the Reading Main Principal Components
Grade 8, Reading to Gain Information

Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect
0	-0.031087	45	-0.039006	90	-0.002636
1	-0.052323	46	0.017230	91	-0.014105
2	-0.005878	47	-0.003005	92	-0.005672
3	-0.078537	48	0.020085	93	0.011805
4	0.123636	49	0.011739	94	-0.016309
5	-0.108430	50	0.013694	95	-0.015879
6	0.024946	51	-0.014678	96	-0.017408
7	-0.046874	52	-0.016132	97	0.002797
8	-0.009914	53	0.015120	98	0.024627
9	0.042651	54	0.013308	99	-0.010684
10	0.097550	55	0.014782	100	-0.013936
11	0.012963	56	-0.029303	101	0.088587
12	-0.020693	57	-0.017598	102	-0.044555
13	-0.000469	58	-0.016942	103	-0.039405
14	-0.031852	59	0.009303	104	-0.029468
15	-0.024762	60	0.007209	105	-0.016461
16	-0.000925	61	0.038701	106	-0.015164
17	-0.043908	62	-0.012930	107	-0.036795
18	-0.010187	63	0.007618	108	-0.005656
19	-0.031583	64	0.004741	109	-0.019919
20	-0.011815	65	-0.011396	110	0.000947
21	0.044395	66	0.001203	111	-0.022534
22	0.000418	67	0.001716		
23	0.027881	68	-0.001700		
24	0.029586	69	0.002079		
25	0.006417	70	-0.005111		
26	0.012211	71	0.005557		
27	-0.022134	72	-0.007011		
28	-0.021006	73	0.011804		
29	-0.012693	74	0.023416		
30	-0.026464	75	-0.000976		
31	-0.002500	76	0.017608		
32	0.020568	77	0.000950		
33	-0.008056	78	-0.017733		
34	-0.004632	79	0.002120		
35	0.006590	80	0.028500		
36	-0.004075	81	0.012911		
37	-0.009334	82	0.024065		
38	0.003964	83	-0.005040		
39	-0.001792	84	0.000681		
40	0.002219	85	-0.007905		
41	-0.008925	86	-0.020512		
42	-0.006558	87	-0.016493		
43	-0.001400	88	-0.000045		
44	0.028688	89	-0.025967		

Table F-13
Estimated Effects for the Reading Main Principal Components
Grade 8, Reading to Perform a Task

Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect
0	-0.369783	45	-0.055986	90	-0.014683
1	-0.087829	46	0.023554	91	-0.024795
2	-0.003047	47	-0.001153	92	-0.002089
3	-0.083681	48	0.039698	93	0.035176
4	0.158365	49	0.004489	94	-0.041313
5	-0.141837	50	0.008183	95	-0.004270
6	0.038845	51	-0.001591	96	0.001068
7	-0.062501	52	-0.002363	97	-0.016482
8	0.001545	53	0.042112	98	0.012756
9	0.039997	54	-0.010705	99	-0.005830
10	0.117281	55	0.003946	100	0.005945
11	0.028804	56	-0.014400	101	0.088119
12	-0.013117	57	-0.035645	102	-0.043264
13	-0.000337	58	-0.023836	103	-0.058852
14	-0.038604	59	0.009667	104	-0.010594
15	-0.041340	60	0.021002	105	-0.019777
16	0.002604	61	0.026151	106	-0.030230
17	-0.043181	62	-0.018455	107	-0.036872
18	0.005539	63	0.017421	108	-0.046638
19	-0.049001	64	0.030156	109	-0.038053
20	-0.017756	65	-0.000630	110	0.015098
21	0.049705	66	0.016728	111	-0.003598
22	-0.000354	67	-0.017442		
23	0.018638	68	0.023647		
24	0.025346	69	-0.014504		
25	0.002743	70	-0.033862		
26	0.016064	71	0.027723		
27	-0.044412	72	0.009052		
28	-0.027793	73	0.005195		
29	-0.020641	74	0.008998		
30	-0.044470	75	-0.000581		
31	0.003819	76	0.021873		
32	0.026619	77	0.028848		
33	-0.029613	78	0.003670		
34	0.002578	79	0.014384		
35	0.011481	80	0.018163		
36	-0.008791	81	0.038514		
37	-0.002234	82	0.022096		
38	-0.001585	83	-0.035568		
39	-0.016867	84	0.001167		
40	-0.010247	85	-0.019186		
41	0.001935	86	-0.008224		
42	0.002617	87	-0.024552		
43	0.001349	88	-0.004882		
44	0.015532	89	-0.034679		

Table F-14
Estimated Effects for the Reading Main Principal Components
Grade 12, Reading for Literary Experience

Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect
0	0.697702	45	-0.036271	90	-0.006923
1	-0.049742	46	0.008920	91	0.006682
2	0.112614	47	0.003787	92	-0.008723
3	-0.013206	48	-0.002438	93	-0.031376
4	0.034475	49	0.015505	94	-0.001844
5	-0.109843	50	0.013223	95	-0.035222
6	-0.056883	51	0.030190	96	-0.049332
7	-0.026900	52	0.008743	97	0.000596
8	-0.039167	53	0.018344	98	-0.011200
9	-0.063198	54	-0.012387	99	-0.025039
10	-0.010808	55	-0.000774	100	-0.034798
11	0.126025	56	0.017751	101	0.000766
12	0.020466	57	-0.026050	102	0.035667
13	-0.088850	58	0.002033	103	0.015594
14	0.024922	59	0.017962	104	-0.024412
15	0.017733	60	-0.006352	105	0.011152
16	0.032139	61	-0.011434	106	-0.027349
17	0.068323	62	-0.016561	107	-0.012434
18	0.020886	63	0.002114	108	0.019588
19	-0.000564	64	-0.030373	109	-0.021540
20	-0.039091	65	-0.028027	110	0.026325
21	-0.002861	66	0.015868	111	0.050370
22	-0.011332	67	-0.021609	112	0.016397
23	0.037138	68	-0.042168	113	0.036941
24	-0.015485	69	0.047676	114	0.056494
25	0.042514	70	-0.007319	115	-0.001401
26	0.009367	71	0.000142		
27	-0.000335	72	-0.016515		
28	-0.041202	73	-0.005471		
29	-0.003303	74	0.002632		
30	-0.021747	75	-0.020125		
31	-0.046818	76	0.015264		
32	-0.005855	77	-0.037694		
33	-0.036503	78	-0.001097		
34	-0.007094	79	-0.002527		
35	-0.000729	80	0.001124		
36	-0.002503	81	0.003381		
37	-0.004826	82	0.042595		
38	-0.008173	83	-0.003562		
39	-0.025875	84	-0.004253		
40	0.000215	85	0.005731		
41	0.016089	86	-0.028255		
42	-0.015263	87	0.011843		
43	0.045780	88	-0.044738		
44	-0.013321	89	0.000719		

Table F-15
Estimated Effects for the Reading Main Principal Components
Grade 12, Reading to Gain Information

Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect
0	0.609125	45	-0.008593	90	-0.017902
1	-0.039424	46	-0.002618	91	0.014904
2	0.090678	47	0.019514	92	-0.000524
3	-0.011689	48	-0.001217	93	-0.015057
4	0.024874	49	0.001542	94	0.008056
5	-0.108640	50	0.006420	95	-0.023538
6	-0.047374	51	0.004914	96	-0.026939
7	-0.011022	52	-0.008426	97	0.006953
8	-0.022335	53	0.013656	98	0.003307
9	-0.044877	54	0.001470	99	-0.005119
10	-0.001284	55	-0.000117	100	-0.011451
11	0.076342	56	0.001413	101	-0.003417
12	0.005400	57	-0.023476	102	0.009049
13	-0.058397	58	-0.000088	103	-0.001229
14	0.031375	59	0.000855	104	-0.034491
15	-0.003788	60	0.005846	105	0.020933
16	0.006464	61	-0.008173	106	-0.017770
17	0.052951	62	-0.001880	107	-0.023338
18	0.027032	63	-0.008348	108	-0.015110
19	-0.008639	64	-0.015408	109	-0.026025
20	-0.018323	65	-0.020149	110	0.035358
21	0.005909	66	0.007759	111	0.013964
22	-0.009953	67	-0.008549	112	0.021151
23	0.008220	68	-0.023267	113	0.030520
24	-0.026751	69	0.014337	114	0.034376
25	0.020240	70	-0.007597	115	0.012327
26	-0.014056	71	-0.003763		
27	-0.004941	72	0.004847		
28	-0.033299	73	0.002876		
29	0.006478	74	0.019567		
30	-0.009660	75	0.003167		
31	-0.028074	76	-0.005791		
32	0.002746	77	-0.021641		
33	-0.022262	78	0.013588		
34	0.009907	79	-0.000767		
35	-0.001772	80	0.001373		
36	-0.004710	81	0.000639		
37	-0.003242	82	0.023611		
38	0.004965	83	-0.000211		
39	-0.013045	84	0.010604		
40	-0.011712	85	-0.001695		
41	0.018169	86	-0.016758		
42	-0.005650	87	0.022529		
43	0.019908	88	-0.014494		
44	-0.012236	89	-0.002065		

Table F-16
Estimated Effects for the Reading Main Principal Components
Grade 12, Reading to Perform a Task

Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect
0	0.353712	45	-0.011911	90	-0.022818
1	-0.048175	46	-0.002712	91	0.008495
2	0.114169	47	-0.001428	92	0.004756
3	-0.017205	48	-0.030207	93	-0.013959
4	0.020673	49	0.015269	94	-0.004928
5	-0.130710	50	-0.007984	95	-0.000293
6	-0.050665	51	0.012281	96	-0.050678
7	-0.011401	52	0.012330	97	0.019510
8	-0.024293	53	0.010339	98	-0.010406
9	-0.055839	54	-0.018269	99	-0.022536
10	0.007577	55	0.019983	100	-0.029513
11	0.098790	56	-0.010838	101	0.004333
12	0.000932	57	-0.034029	102	0.015110
13	-0.069157	58	-0.007206	103	0.054092
14	0.040781	59	0.027551	104	-0.010099
15	0.013087	60	-0.001890	105	0.014051
16	0.010901	61	-0.011555	106	0.006631
17	0.079433	62	0.001311	107	-0.045771
18	0.036499	63	0.008570	108	-0.015993
19	-0.004586	64	-0.024943	109	-0.028107
20	-0.022112	65	-0.025399	110	0.029893
21	-0.006567	66	0.025859	111	0.022587
22	-0.027376	67	-0.051492	112	0.032881
23	0.028056	68	-0.005005	113	0.056744
24	-0.010786	69	0.015431	114	0.054458
25	0.019892	70	0.009834	115	0.010551
26	0.009480	71	-0.015615		
27	-0.020742	72	0.006093		
28	-0.043454	73	-0.011430		
29	0.021287	74	0.007885		
30	-0.011662	75	0.003677		
31	-0.001434	76	-0.017501		
32	0.002993	77	-0.014745		
33	-0.011036	78	0.020172		
34	-0.002046	79	0.000804		
35	-0.016133	80	-0.011243		
36	-0.013061	81	-0.007945		
37	-0.021196	82	0.019390		
38	0.003525	83	-0.006550		
39	0.014867	84	0.009335		
40	-0.002993	85	0.026388		
41	0.007958	86	-0.017412		
42	-0.025246	87	-0.002386		
43	0.000013	88	-0.024869		
44	-0.001626	89	0.000211		

Table F-17
Estimated Effects for the Mathematics Main Principal Components
Grade 4, Numbers and Operations

Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect
0	-0.007747	45	0.030606	90	0.000424	135	0.006252
1	-0.038809	46	-0.006707	91	-0.000341	136	0.019045
2	-0.023154	47	0.002517	92	-0.019645	137	-0.020008
3	0.237634	48	-0.028434	93	-0.006482	138	-0.027683
4	-0.059482	49	0.003768	94	-0.035726	139	0.009777
5	-0.021540	50	0.039163	95	0.030275	140	-0.030077
6	-0.029968	51	-0.008883	96	-0.027050	141	0.040908
7	0.008610	52	-0.019066	97	0.011197	142	0.012034
8	-0.053139	53	0.009155	98	0.007983	143	-0.008004
9	0.071217	54	-0.003730	99	0.041621		
10	-0.071793	55	-0.007937	100	-0.043569		
11	0.019322	56	-0.007244	101	-0.008097		
12	0.000609	57	-0.012942	102	-0.046143		
13	0.019913	58	0.002364	103	-0.025765		
14	-0.018794	59	-0.002056	104	0.029040		
15	-0.008113	60	0.012563	105	0.025773		
16	0.004333	61	-0.029445	106	0.002471		
17	0.020894	62	0.009910	107	0.030239		
18	0.032789	63	0.001110	108	0.010515		
19	-0.010737	64	-0.018918	109	-0.004180		
20	-0.000007	65	0.020152	110	0.005534		
21	-0.014162	66	0.014548	111	0.004321		
22	0.010161	67	-0.015463	112	0.033383		
23	0.030900	68	0.004604	113	-0.044625		
24	0.018845	69	-0.022477	114	-0.028843		
25	-0.009518	70	0.007523	115	0.006748		
26	0.011442	71	-0.036888	116	-0.001835		
27	0.001814	72	-0.034687	117	-0.037833		
28	-0.095046	73	-0.011177	118	-0.005145		
29	-0.019956	74	-0.017237	119	-0.019477		
30	0.004234	75	-0.006555	120	-0.041876		
31	0.002388	76	0.026702	121	-0.018151		
32	0.037028	77	-0.011639	122	0.004883		
33	0.006273	78	-0.041342	123	-0.027520		
34	-0.045301	79	0.016502	124	-0.022503		
35	-0.025501	80	-0.014027	125	-0.055371		
36	-0.001790	81	-0.060746	126	0.005752		
37	-0.026215	82	-0.005165	127	0.046626		
38	0.070247	83	-0.018997	128	0.036665		
39	0.071440	84	0.000481	129	0.059566		
40	0.021646	85	-0.000121	130	0.005848		
41	0.049507	86	0.042929	131	-0.027680		
42	-0.036765	87	0.014766	132	0.028441		
43	-0.042430	88	-0.012685	133	-0.011283		
44	-0.003410	89	0.030329	134	0.027264		

Table F-18
Estimated Effects for the Mathematics Main Principal Components
Grade 4, Measurement

Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect
0	0.002345	45	0.026890	90	-0.004997	135	-0.001714
1	-0.038764	46	-0.014065	91	-0.018659	136	-0.008969
2	-0.026576	47	-0.003256	92	0.012859	137	0.000941
3	0.250827	48	-0.026186	93	0.001605	138	-0.018633
4	-0.040026	49	0.015980	94	-0.039096	139	0.020222
5	-0.002567	50	0.032971	95	0.025490	140	-0.028542
6	-0.017750	51	0.001117	96	-0.024771	141	0.017412
7	0.036742	52	-0.021427	97	0.021943	142	-0.006357
8	-0.055342	53	-0.015765	98	-0.003978	143	0.018454
9	0.035036	54	0.005368	99	0.046750		
10	-0.066886	55	0.000641	100	-0.041652		
11	0.014383	56	-0.020666	101	0.006076		
12	0.004844	57	0.002928	102	-0.035623		
13	0.014365	58	-0.000037	103	-0.011605		
14	-0.001696	59	-0.021834	104	0.026547		
15	-0.001260	60	-0.007435	105	0.019443		
16	0.009970	61	-0.049216	106	-0.024051		
17	0.004237	62	0.007424	107	0.010547		
18	0.036040	63	0.017486	108	-0.002730		
19	-0.027886	64	-0.035796	109	-0.015716		
20	-0.004414	65	0.015541	110	0.021430		
21	-0.015915	66	0.042170	111	0.023540		
22	0.017383	67	-0.003346	112	0.029576		
23	0.044602	68	0.006114	113	-0.006028		
24	0.017105	69	-0.019620	114	-0.034692		
25	0.020381	70	0.013473	115	-0.000134		
26	0.014326	71	-0.051372	116	-0.015903		
27	-0.015755	72	-0.036153	117	-0.037614		
28	-0.083693	73	-0.009593	118	0.032212		
29	-0.003177	74	-0.006280	119	-0.029269		
30	-0.016423	75	0.005028	120	-0.010367		
31	-0.007840	76	0.040615	121	-0.005922		
32	0.014575	77	-0.015205	122	0.013453		
33	0.014166	78	-0.022144	123	-0.011189		
34	-0.037992	79	0.004012	124	-0.009962		
35	-0.019061	80	-0.006763	125	-0.052952		
36	0.003680	81	-0.077474	126	0.015732		
37	-0.025768	82	-0.018444	127	0.046054		
38	0.070810	83	-0.012455	128	0.048721		
39	0.045005	84	0.011185	129	0.060481		
40	0.006608	85	-0.010708	130	-0.012885		
41	0.062261	86	0.050011	131	-0.043281		
42	-0.028271	87	0.010637	132	0.014527		
43	-0.048086	88	-0.018096	133	-0.037571		
44	0.000163	89	0.037927	134	0.008549		

Table F-19
Estimated Effects for the Mathematics Main Principal Components
Grade 4, Geometry

Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect
0	0.004741	45	0.014894	90	-0.018508	135	0.005079
1	-0.035401	46	0.004156	91	-0.022792	136	0.046386
2	-0.032017	47	0.010669	92	0.025738	137	-0.013458
3	0.232311	48	-0.040897	93	-0.011446	138	-0.010600
4	-0.050377	49	0.011130	94	-0.021855	139	-0.017825
5	-0.002275	50	0.038320	95	0.018594	140	0.002875
6	-0.005345	51	-0.027600	96	-0.011781	141	0.006092
7	0.042659	52	-0.010982	97	0.007763	142	0.003942
8	-0.077495	53	0.018805	98	0.029739	143	0.023333
9	0.047893	54	0.007967	99	0.063990		
10	-0.036399	55	-0.012891	100	-0.038657		
11	0.026388	56	-0.034381	101	0.008596		
12	-0.007501	57	0.023866	102	-0.051031		
13	0.013215	58	0.001262	103	-0.003895		
14	-0.002823	59	-0.017458	104	0.008592		
15	-0.024276	60	0.011196	105	0.018511		
16	0.007592	61	-0.059010	106	-0.021504		
17	-0.002922	62	0.005217	107	-0.005347		
18	0.009899	63	-0.000430	108	0.023470		
19	-0.003682	64	-0.019316	109	-0.009992		
20	-0.011565	65	-0.019308	110	-0.030644		
21	-0.018199	66	0.048537	111	0.022078		
22	-0.001977	67	0.017442	112	0.040148		
23	0.028313	68	0.012638	113	-0.017020		
24	0.030481	69	-0.007491	114	-0.047599		
25	0.022853	70	0.009310	115	0.012231		
26	0.000455	71	-0.044609	116	0.012714		
27	-0.009940	72	-0.022548	117	-0.051808		
28	-0.067498	73	-0.007572	118	0.024463		
29	0.000302	74	-0.006807	119	-0.028571		
30	-0.005063	75	-0.014558	120	-0.009414		
31	0.007302	76	0.029971	121	-0.018808		
32	0.026429	77	-0.035213	122	0.020938		
33	0.007817	78	-0.024530	123	-0.022669		
34	-0.060932	79	0.033701	124	-0.010582		
35	-0.025840	80	-0.018124	125	-0.038664		
36	0.035182	81	-0.044643	126	0.013878		
37	-0.017514	82	-0.010975	127	0.018802		
38	0.052463	83	-0.016432	128	0.020052		
39	0.069096	84	0.022601	129	0.067230		
40	0.000636	85	-0.000179	130	0.009068		
41	0.054831	86	0.044966	131	-0.030649		
42	-0.024213	87	0.021078	132	0.041964		
43	-0.044840	88	0.001339	133	-0.038247		
44	0.010719	89	0.016578	134	0.030475		

Table F-20
Estimated Effects for the Mathematics Main Principal Components
Grade 4, Data Analysis, Statistics, and Probability

Principal Component	Estimated Effect
0	-0.039669
1	-0.042420
2	-0.034575
3	0.251205
4	-0.055593
5	-0.021174
6	-0.028804
7	0.020695
8	-0.051515
9	0.070102
10	-0.071113
11	0.012250
12	-0.002875
13	0.028311
14	-0.006099
15	-0.005986
16	-0.004757
17	0.000459
18	0.026010
19	-0.001277
20	0.003746
21	-0.022758
22	0.001116
23	0.028614
24	0.030040
25	0.001229
26	0.039123
27	-0.007471
28	-0.083293
29	-0.004164
30	-0.004304
31	0.000988
32	0.040016
33	0.010659
34	-0.049852
35	-0.032465
36	0.014218
37	-0.015030
38	0.061736
39	0.060860
40	0.012007
41	0.053509
42	-0.036504
43	-0.041133
44	-0.011510
45	0.033425
46	-0.000518
47	-0.006415
48	-0.029899
49	0.004381
50	0.055447
51	-0.021818
52	-0.025459
53	-0.000118
54	-0.012163
55	0.010970
56	-0.011899
57	0.000178
58	0.001227
59	-0.020309
60	0.009282
61	-0.070111
62	0.021735
63	0.004845
64	-0.035165
65	0.022302
66	0.018146
67	-0.010619
68	0.011304
69	-0.018762
70	0.003670
71	-0.060911
72	-0.037276
73	-0.033906
74	-0.007386
75	-0.005318
76	0.022653
77	-0.022754
78	-0.031452
79	0.010866
80	-0.036012
81	-0.080255
82	-0.009881
83	-0.011982
84	0.004383
85	0.010562
86	0.043289
87	-0.003346
88	-0.027372
89	0.020999
90	0.018751
91	-0.014809
92	-0.015889
93	0.008226
94	-0.041767
95	0.033789
96	-0.001240
97	0.016195
98	0.000560
99	0.054643
100	-0.013712
101	-0.015487
102	-0.045733
103	-0.027831
104	0.007373
105	0.028225
106	-0.001420
107	0.033498
108	0.043523
109	0.000223
110	-0.017048
111	0.000159
112	0.037077
113	-0.012823
114	-0.030547
115	0.001643
116	0.002450
117	-0.050207
118	0.007141
119	-0.040200
120	-0.030847
121	-0.038863
122	0.001575
123	-0.052422
124	-0.016665
125	-0.050145
126	0.004102
127	0.047125
128	0.034725
129	0.068055
130	0.000624
131	-0.038372
132	0.022466
133	-0.003150
134	0.038229
135	0.017357
136	0.011290
137	-0.019916
138	-0.004772
139	0.019403
140	-0.033746
141	0.020589
142	-0.006206
143	0.005397

Table F-21
Estimated Effects for the Mathematics Main Principal Components
Grade 4, Algebra and Functions

Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect
0	-0.081408	45	0.005828	90	0.000183	135	-0.000362
1	-0.039302	46	-0.005016	91	-0.023720	136	0.010702
2	-0.034880	47	0.012705	92	-0.014315	137	0.023562
3	0.263793	48	-0.023738	93	-0.006807	138	-0.057633
4	-0.066680	49	0.002068	94	-0.025348	139	0.003370
5	-0.034291	50	0.063155	95	0.014550	140	-0.058643
6	-0.028426	51	-0.032130	96	-0.003568	141	0.019656
7	0.027750	52	0.003428	97	0.022114	142	0.011606
8	-0.062779	53	0.015088	98	0.022724	143	0.039308
9	0.050144	54	-0.019500	99	0.037276		
10	-0.067467	55	-0.014391	100	-0.005996		
11	0.008914	56	-0.036127	101	0.005090		
12	-0.027427	57	0.003066	102	-0.063825		
13	0.024687	58	-0.021252	103	-0.016181		
14	-0.007056	59	-0.004621	104	-0.006123		
15	-0.011700	60	0.015394	105	0.018266		
16	0.003195	61	-0.064528	106	0.010720		
17	0.009805	62	0.020381	107	0.042541		
18	0.037124	63	0.006481	108	0.047117		
19	-0.020193	64	-0.033061	109	-0.021765		
20	-0.016137	65	-0.001241	110	0.011014		
21	-0.006946	66	0.031945	111	0.017647		
22	0.008594	67	-0.020384	112	0.064799		
23	0.023399	68	0.007911	113	-0.024515		
24	0.042468	69	-0.049267	114	-0.054932		
25	0.004148	70	0.005285	115	0.007126		
26	0.010288	71	-0.066334	116	-0.002981		
27	0.005778	72	-0.018584	117	-0.039291		
28	-0.088904	73	-0.018256	118	0.018395		
29	-0.034327	74	0.027600	119	-0.023785		
30	0.002398	75	-0.014078	120	-0.023376		
31	0.009072	76	0.021033	121	-0.049626		
32	0.035275	77	-0.030003	122	0.016192		
33	0.022282	78	-0.055088	123	-0.046964		
34	-0.022867	79	0.045671	124	-0.036659		
35	-0.038001	80	-0.013966	125	-0.044010		
36	0.012601	81	-0.067823	126	0.007700		
37	-0.013418	82	-0.033280	127	0.045437		
38	0.061235	83	-0.041746	128	0.054637		
39	0.051386	84	-0.000427	129	0.068208		
40	0.004712	85	0.010267	130	0.042000		
41	0.068398	86	0.054419	131	-0.065688		
42	-0.030704	87	0.003290	132	0.048200		
43	-0.042140	88	-0.027001	133	-0.049819		
44	-0.009301	89	0.023614	134	0.024538		

Table F-22
Estimated Effects for the Mathematics Main Principal Components
Grade 8, Numbers and Operations

Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect
0	0.030977	45	0.009560	90	-0.010469	135	-0.003355
1	-0.034988	46	0.001702	91	-0.029376	136	-0.026263
2	-0.098747	47	0.019224	92	0.028028	137	-0.001980
3	-0.084890	48	-0.018550	93	-0.036721	138	-0.018428
4	-0.190847	49	0.034962	94	0.017034	139	-0.016813
5	-0.054843	50	-0.013986	95	-0.015974	140	0.037776
6	0.045479	51	0.002016	96	-0.030783	141	-0.029225
7	-0.018873	52	0.035985	97	0.021760	142	0.026234
8	0.063847	53	0.042183	98	0.016181	143	0.013749
9	-0.029807	54	-0.017096	99	0.007475	144	0.059361
10	-0.027829	55	0.017040	100	0.017603	145	-0.014533
11	-0.039737	56	0.014109	101	-0.006151	146	0.027289
12	0.111929	57	-0.009611	102	-0.018129	147	-0.010967
13	-0.002385	58	-0.028189	103	0.047037	148	-0.016582
14	0.067753	59	-0.022683	104	-0.001250	149	-0.009336
15	0.034703	60	-0.018130	105	-0.013596	150	-0.002152
16	0.033877	61	-0.016100	106	-0.001817	151	0.008391
17	0.009208	62	-0.011063	107	-0.019918	152	-0.009537
18	-0.008366	63	-0.011434	108	0.013037	153	0.015544
19	-0.002794	64	0.011647	109	-0.024628	154	-0.018611
20	-0.035046	65	0.014236	110	-0.025904	155	-0.003647
21	0.001113	66	0.030809	111	0.027152	156	0.011856
22	0.013292	67	0.000489	112	0.004614	157	0.035663
23	-0.038429	68	0.023627	113	-0.000283	158	0.004090
24	0.006202	69	-0.033867	114	-0.021400	159	0.018862
25	0.009534	70	-0.006573	115	-0.028477	160	0.015649
26	-0.043835	71	0.018421	116	-0.000665		
27	-0.013388	72	-0.008019	117	-0.019739		
28	-0.017398	73	-0.001339	118	0.011885		
29	-0.003682	74	0.009518	119	-0.012363		
30	0.037822	75	0.003166	120	0.013522		
31	0.011000	76	-0.021239	121	0.029096		
32	-0.066980	77	0.001415	122	0.013692		
33	0.000763	78	-0.006774	123	-0.016986		
34	-0.040187	79	-0.012937	124	-0.028156		
35	0.002539	80	-0.004937	125	-0.005387		
36	0.000616	81	-0.016521	126	0.026344		
37	0.026097	82	0.003422	127	0.000618		
38	-0.011172	83	-0.014122	128	-0.008961		
39	-0.011537	84	-0.006225	129	0.039067		
40	0.002458	85	0.014279	130	0.003468		
41	0.019567	86	0.000167	131	0.003684		
42	0.007098	87	-0.015414	132	-0.002371		
43	0.053684	88	-0.021400	133	-0.007848		
44	0.005010	89	-0.002948	134	-0.002661		

Table F-23
Estimated Effects for the Mathematics Main Principal Components
Grade 8, Measurement

Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect
0	0.028429	45	0.022022	90	0.004459	135	0.000389
1	-0.032352	46	0.003320	91	-0.033267	136	-0.005871
2	-0.102854	47	0.029826	92	0.021553	137	-0.035580
3	-0.087899	48	-0.025065	93	-0.040770	138	0.009462
4	-0.198460	49	0.026220	94	-0.004106	139	-0.043921
5	-0.052836	50	-0.016938	95	0.000990	140	0.038785
6	0.018604	51	0.006521	96	-0.036534	141	-0.012250
7	-0.047133	52	0.014610	97	0.013166	142	0.025530
8	0.037762	53	0.044633	98	-0.002316	143	0.006753
9	-0.039394	54	-0.027735	99	0.005759	144	0.062892
10	-0.045022	55	0.029795	100	0.004539	145	0.000185
11	-0.057245	56	0.003949	101	0.012207	146	0.034334
12	0.099751	57	-0.033466	102	0.008833	147	0.011683
13	-0.018863	58	-0.015158	103	0.022204	148	-0.039295
14	0.080366	59	-0.019466	104	0.025552	149	0.004877
15	0.049333	60	0.016086	105	-0.017520	150	-0.008654
16	0.025899	61	-0.004424	106	0.013399	151	-0.007006
17	0.008292	62	-0.009475	107	-0.010065	152	-0.024986
18	-0.005343	63	-0.001443	108	0.001541	153	0.029975
19	0.007024	64	0.021171	109	-0.017224	154	-0.042001
20	-0.016906	65	-0.000002	110	-0.036754	155	-0.010994
21	0.001366	66	0.030129	111	0.008722	156	-0.001441
22	0.034540	67	0.005197	112	-0.004902	157	0.040560
23	-0.040656	68	0.023332	113	-0.007878	158	0.001804
24	0.017870	69	-0.041995	114	0.004118	159	0.033608
25	0.015303	70	0.004478	115	-0.010314	160	0.006148
26	-0.045610	71	0.016582	116	0.009446		
27	-0.002206	72	0.020114	117	-0.033280		
28	-0.026738	73	-0.021960	118	0.030105		
29	-0.012152	74	0.007223	119	-0.019346		
30	0.056124	75	-0.000055	120	0.006060		
31	0.005069	76	-0.036078	121	0.020580		
32	-0.071202	77	0.004811	122	0.026114		
33	0.020975	78	-0.008476	123	-0.027480		
34	-0.004028	79	-0.026531	124	-0.067431		
35	0.016695	80	0.004100	125	-0.036897		
36	-0.012036	81	-0.000936	126	0.034397		
37	0.013372	82	-0.012357	127	-0.009302		
38	-0.020658	83	0.005786	128	0.000415		
39	-0.002843	84	-0.016614	129	0.024500		
40	0.008515	85	0.013884	130	-0.016681		
41	0.036210	86	-0.025346	131	0.015433		
42	-0.003662	87	-0.012877	132	0.003812		
43	0.056058	88	-0.019004	133	-0.030635		
44	-0.015499	89	-0.002817	134	-0.003436		

Table F-24
Estimated Effects for the Mathematics Main Principal Components
Grade 8, Geometry

Principal Component	Estimated Effect
0	0.005772
1	-0.029816
2	-0.099621
3	-0.081329
4	-0.184633
5	-0.053703
6	0.027639
7	-0.049665
8	0.035994
9	-0.028785
10	-0.037620
11	-0.045289
12	0.089394
13	-0.004920
14	0.072368
15	0.041712
16	0.012681
17	0.005066
18	0.001830
19	0.003689
20	-0.029271
21	-0.013529
22	0.009124
23	-0.034080
24	0.009035
25	0.004293
26	-0.040156
27	-0.019361
28	-0.020952
29	-0.022882
30	0.042489
31	0.002596
32	-0.062309
33	0.016637
34	-0.016309
35	0.004781
36	0.008176
37	0.003524
38	-0.002138
39	-0.015345
40	0.002342
41	0.030321
42	0.016314
43	0.045713
44	-0.004291
45	0.004909
46	0.006147
47	0.018342
48	-0.054199
49	0.042907
50	-0.036324
51	-0.013978
52	0.021179
53	0.040934
54	-0.024216
55	0.014555
56	0.002607
57	-0.016966
58	-0.018180
59	-0.020439
60	0.012961
61	0.002769
62	-0.005786
63	-0.006673
64	0.015148
65	-0.001195
66	0.024356
67	-0.002843
68	0.008563
69	-0.040374
70	0.006471
71	0.005155
72	0.021764
73	-0.019200
74	0.015612
75	-0.010606
76	-0.030122
77	-0.003370
78	0.003890
79	-0.016720
80	0.011422
81	-0.003671
82	-0.011656
83	-0.016649
84	0.011329
85	0.019188
86	-0.016024
87	-0.019919
88	-0.021408
89	0.001535
90	-0.000652
91	-0.033459
92	0.008847
93	-0.041478
94	-0.017466
95	0.003255
96	-0.026525
97	0.014567
98	0.010420
99	0.001857
100	0.008294
101	-0.006302
102	-0.003363
103	0.044847
104	-0.001417
105	-0.018745
106	-0.014436
107	-0.009745
108	-0.027085
109	-0.015797
110	-0.024221
111	0.011503
112	-0.002057
113	-0.000922
114	-0.015836
115	-0.012477
116	0.016307
117	-0.032596
118	0.016823
119	-0.033480
120	0.002531
121	0.034877
122	0.020022
123	-0.037381
124	-0.034782
125	-0.029208
126	0.049898
127	-0.015547
128	-0.015511
129	0.037820
130	0.017624
131	0.011196
132	0.003036
133	-0.024016
134	0.010171
135	0.001759
136	-0.000704
137	0.005852
138	-0.006801
139	-0.019170
140	0.008898
141	-0.032685
142	0.026559
143	0.010363
144	0.049900
145	-0.007097
146	0.048883
147	-0.009022
148	-0.019800
149	0.002393
150	-0.026056
151	0.003620
152	0.003455
153	0.028097
154	-0.048892
155	0.014265
156	0.014910
157	0.030082
158	-0.001053
159	0.005942
160	0.007882

Table F-25
Estimated Effects for the Mathematics Main Principal Components
Grade 8, Data Analysis, Statistics, and Probability

Principal Component	Estimated Effect
0	0.028219
1	-0.034023
2	-0.099200
3	-0.082233
4	-0.197984
5	-0.068331
6	0.033238
7	-0.033246
8	0.049514
9	-0.031074
10	-0.027089
11	-0.057641
12	0.094418
13	-0.015545
14	0.063658
15	0.028903
16	0.038278
17	0.003118
18	-0.012341
19	-0.005524
20	-0.025747
21	0.006934
22	0.007119
23	-0.030575
24	0.006645
25	-0.007408
26	-0.063489
27	-0.026103
28	-0.020619
29	-0.012637
30	0.028881
31	-0.006204
32	-0.071989
33	-0.004221
34	-0.024268
35	0.009123
36	0.004055
37	0.028007
38	-0.009910
39	-0.012688
40	-0.001002
41	0.020901
42	0.009740
43	0.032459
44	-0.004427
45	0.020566
46	0.022029
47	0.013306
48	-0.033101
49	0.018431
50	-0.008525
51	0.011387
52	0.045869
53	0.042869
54	-0.025769
55	0.026537
56	0.006125
57	-0.004626
58	-0.006846
59	-0.022950
60	-0.000031
61	-0.014450
62	0.006563
63	-0.013716
64	0.015577
65	0.022839
66	0.035755
67	0.011451
68	0.001582
69	-0.038906
70	-0.021955
71	0.005839
72	-0.014618
73	0.000641
74	0.005936
75	0.000388
76	-0.019111
77	0.003252
78	-0.011967
79	-0.007335
80	0.001587
81	-0.018078
82	-0.010896
83	0.006977
84	-0.015605
85	0.008411
86	-0.001372
87	-0.035877
88	-0.007404
89	-0.002157
90	0.005240
91	-0.019463
92	0.034825
93	-0.043411
94	0.001233
95	-0.003860
96	-0.029578
97	0.006641
98	0.000359
99	0.005597
100	0.000885
101	-0.021783
102	-0.001645
103	0.049367
104	0.000393
105	-0.026845
106	-0.015119
107	-0.028106
108	-0.013234
109	-0.008002
110	-0.024164
111	0.017370
112	0.006094
113	0.015452
114	-0.019036
115	-0.011957
116	0.008307
117	-0.014099
118	0.001685
119	-0.040962
120	0.003870
121	0.022659
122	0.012421
123	-0.005110
124	-0.031410
125	-0.018904
126	0.017520
127	-0.004201
128	-0.012774
129	0.044753
130	0.001035
131	0.002659
132	0.005824
133	-0.022956
134	-0.023039
135	0.009886
136	-0.013126
137	0.004481
138	-0.001513
139	0.005740
140	0.020681
141	-0.037398
142	0.026036
143	0.000758
144	0.052432
145	-0.010648
146	0.034217
147	-0.027589
148	-0.024010
149	-0.006082
150	0.006606
151	0.002128
152	-0.014496
153	0.043536
154	-0.019221
155	0.001706
156	0.016762
157	0.034107
158	-0.006932
159	0.014163
160	0.014716

Table F-26
Estimated Effects for the Mathematics Main Principal Components
Grade 8, Algebra and Functions

Principal Component	Estimated Effect
0	0.030777
1	-0.036733
2	-0.101231
3	-0.079000
4	-0.195796
5	-0.051868
6	0.055071
7	-0.034530
8	0.072078
9	-0.033210
10	-0.045254
11	-0.051384
12	0.104292
13	0.014041
14	0.067236
15	0.047226
16	0.017458
17	0.000504
18	-0.005420
19	-0.005497
20	-0.010204
21	-0.013422
22	0.006427
23	-0.040827
24	0.000522
25	0.009157
26	-0.049100
27	-0.011387
28	-0.021500
29	-0.015206
30	0.041705
31	0.000623
32	-0.058661
33	0.000997
34	-0.018508
35	-0.000821
36	0.011680
37	0.019825
38	0.000122
39	-0.003999
40	-0.018122
41	0.022521
42	0.004858
43	0.061955
44	-0.001843
45	0.021489
46	0.012479
47	0.005897
48	-0.056568
49	0.012230
50	-0.032036
51	0.007164
52	0.033298
53	0.050760
54	-0.019349
55	0.025423
56	-0.004087
57	-0.019291
58	-0.018223
59	-0.008832
60	0.006051
61	-0.021149
62	-0.021834
63	-0.003866
64	0.017450
65	0.025885
66	0.019609
67	0.007306
68	-0.001136
69	-0.034519
70	0.007878
71	-0.008084
72	0.002724
73	0.003265
74	-0.002070
75	-0.003857
76	-0.023364
77	0.008610
78	-0.003758
79	0.005888
80	-0.005855
81	-0.009788
82	-0.004705
83	-0.025089
84	0.001219
85	0.017422
86	0.000198
87	-0.027881
88	-0.025192
89	-0.002082
90	-0.004785
91	-0.023697
92	0.035903
93	-0.049708
94	0.007051
95	-0.033791
96	-0.024256
97	0.003673
98	0.011416
99	0.010270
100	0.011012
101	-0.004797
102	-0.014998
103	0.048558
104	-0.010109
105	-0.032417
106	-0.019002
107	-0.015294
108	-0.024185
109	-0.013028
110	-0.040320
111	0.030839
112	-0.017334
113	0.005526
114	-0.032580
115	-0.021323
116	-0.000942
117	-0.024040
118	0.030428
119	-0.020534
120	0.016717
121	0.027139
122	0.009256
123	-0.005858
124	-0.050844
125	-0.039143
126	0.028341
127	-0.006973
128	0.003291
129	0.050258
130	0.012890
131	-0.003707
132	0.002454
133	-0.013105
134	0.001156
135	0.006318
136	-0.013588
137	0.006452
138	0.000617
139	0.008016
140	0.011464
141	-0.041295
142	0.026494
143	0.029759
144	0.058440
145	-0.020456
146	0.022348
147	0.000091
148	-0.026506
149	-0.011300
150	-0.010487
151	0.017158
152	-0.000462
153	0.000247
154	-0.017019
155	0.003417
156	0.027273
157	0.036739
158	-0.005852
159	0.011069
160	0.002697

Table F-27
Estimated Effects for the Mathematics Main Principal Components
Grade 12, Numbers and Operations

Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect
0	0.039152	45	0.017843	90	0.050014	135	0.014608
1	0.167601	46	0.023440	91	-0.009334	136	-0.008615
2	0.009960	47	0.028183	92	-0.025635	137	-0.043001
3	0.005775	48	-0.004528	93	0.035172	138	-0.038446
4	-0.059097	49	0.015832	94	0.021429		
5	-0.071251	50	-0.023527	95	-0.012814		
6	0.070414	51	-0.031408	96	-0.023667		
7	0.030412	52	-0.006749	97	0.001590		
8	-0.075242	53	0.014096	98	0.014136		
9	0.003044	54	-0.017178	99	0.009408		
10	-0.023873	55	0.003365	100	-0.021778		
11	-0.028261	56	0.007345	101	0.004756		
12	-0.099318	57	-0.014769	102	-0.037187		
13	0.009766	58	-0.015585	103	0.036994		
14	0.048835	59	-0.016405	104	-0.027974		
15	-0.015697	60	-0.007533	105	0.000275		
16	-0.053647	61	-0.003126	106	-0.003402		
17	0.013714	62	0.008774	107	-0.016949		
18	-0.032559	63	-0.002751	108	-0.012644		
19	-0.039348	64	0.002782	109	-0.005159		
20	0.028105	65	0.029552	110	0.009027		
21	-0.018901	66	0.013281	111	0.000176		
22	0.013572	67	0.017837	112	0.003154		
23	0.054234	68	-0.000829	113	0.005074		
24	-0.014328	69	-0.005167	114	-0.010049		
25	-0.000945	70	-0.009942	115	0.011666		
26	0.004834	71	0.003097	116	-0.053227		
27	-0.003530	72	0.025774	117	-0.005474		
28	0.005991	73	-0.034964	118	-0.013893		
29	-0.016823	74	0.012754	119	0.002568		
30	0.003089	75	0.005141	120	-0.005726		
31	-0.032731	76	0.000849	121	0.016011		
32	-0.027150	77	-0.017078	122	0.032523		
33	0.003379	78	-0.016606	123	0.010943		
34	0.009164	79	-0.004654	124	-0.028451		
35	0.024102	80	-0.015437	125	-0.062373		
36	-0.002547	81	-0.012962	126	-0.007252		
37	0.012774	82	0.008834	127	0.028083		
38	0.018611	83	0.016801	128	0.026536		
39	0.035952	84	-0.010346	129	-0.003056		
40	-0.016857	85	-0.001785	130	0.028312		
41	0.014501	86	-0.032770	131	-0.016060		
42	-0.006935	87	-0.010835	132	-0.001673		
43	-0.013438	88	0.027462	133	-0.030841		
44	-0.004391	89	-0.007012	134	0.020850		

Table F-28
Estimated Effects for the Mathematics Main Principal Components
Grade 12, Measurement

Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect
0	0.042479	45	0.025980	90	0.056498	135	0.032091
1	0.168985	46	0.028298	91	0.005371	136	-0.035301
2	0.009600	47	0.012162	92	-0.022258	137	-0.055750
3	0.009854	48	0.002445	93	0.022745	138	-0.041784
4	-0.066495	49	-0.004161	94	0.012602		
5	-0.082867	50	-0.016924	95	-0.002059		
6	0.067655	51	-0.040921	96	-0.013718		
7	0.003218	52	-0.015554	97	0.011478		
8	-0.074674	53	-0.016993	98	0.003774		
9	0.033729	54	-0.023018	99	-0.013750		
10	-0.030317	55	0.009241	100	-0.026281		
11	-0.042789	56	0.003727	101	0.004143		
12	-0.120540	57	-0.005391	102	-0.027121		
13	-0.019190	58	-0.007735	103	0.020119		
14	0.035925	59	-0.017678	104	-0.040703		
15	0.005915	60	-0.007916	105	0.018575		
16	-0.039622	61	-0.030694	106	0.002170		
17	0.019910	62	0.008661	107	0.004815		
18	-0.012054	63	0.013006	108	-0.046322		
19	-0.032975	64	0.011337	109	-0.004883		
20	0.014006	65	0.016475	110	-0.008730		
21	-0.022367	66	-0.008412	111	0.006969		
22	0.022303	67	0.024053	112	0.010078		
23	0.034567	68	0.002446	113	-0.001559		
24	-0.008431	69	-0.004471	114	-0.020967		
25	0.003254	70	-0.004744	115	-0.005111		
26	0.027024	71	0.022162	116	-0.058728		
27	-0.005424	72	0.027340	117	-0.023020		
28	0.022143	73	-0.018634	118	-0.001371		
29	-0.020867	74	0.016757	119	0.031867		
30	-0.009811	75	-0.000232	120	0.008447		
31	-0.019038	76	-0.040666	121	0.017623		
32	-0.015752	77	-0.007699	122	0.020319		
33	0.013125	78	-0.013198	123	0.014011		
34	0.009441	79	-0.017518	124	-0.010477		
35	0.024070	80	-0.010895	125	-0.062689		
36	-0.009077	81	-0.015203	126	-0.006826		
37	0.029124	82	0.000673	127	0.011744		
38	0.015159	83	0.016223	128	-0.002209		
39	0.018217	84	-0.005290	129	-0.010303		
40	-0.007385	85	-0.016358	130	0.046291		
41	0.019395	86	-0.028111	131	-0.005996		
42	0.010922	87	-0.008398	132	0.006076		
43	-0.008707	88	0.037445	133	-0.051140		
44	0.012200	89	-0.007543	134	0.055407		

Table F-29
Estimated Effects for the Mathematics Main Principal Components
Grade 12, Geometry

Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect
0	0.027732	45	0.027343	90	0.040909	135	0.002564
1	0.169569	46	0.035672	91	-0.007378	136	-0.036736
2	0.013552	47	0.030061	92	-0.021916	137	-0.047602
3	0.011660	48	0.014364	93	0.025275	138	-0.019826
4	-0.087342	49	0.005790	94	0.000444		
5	-0.075516	50	-0.017185	95	0.006629		
6	0.072843	51	-0.001287	96	0.002254		
7	0.037955	52	0.005530	97	0.000186		
8	-0.067149	53	-0.017246	98	-0.010987		
9	0.002460	54	-0.022510	99	-0.005400		
10	-0.027035	55	-0.007953	100	-0.016364		
11	-0.051958	56	0.005310	101	-0.009095		
12	-0.100042	57	0.002654	102	-0.027664		
13	-0.007574	58	0.007577	103	0.035879		
14	0.024001	59	-0.014663	104	-0.034901		
15	0.006113	60	0.011279	105	0.008528		
16	-0.027388	61	-0.032602	106	-0.003332		
17	0.021945	62	0.016237	107	0.003456		
18	-0.003519	63	0.012468	108	-0.031922		
19	-0.037425	64	0.019497	109	-0.004495		
20	-0.002158	65	0.013277	110	0.013289		
21	-0.015513	66	-0.006285	111	0.006107		
22	0.030414	67	0.018770	112	0.007423		
23	0.005188	68	0.007207	113	-0.001957		
24	-0.006454	69	-0.008287	114	-0.034009		
25	0.008097	70	0.001292	115	0.013774		
26	0.004379	71	-0.000931	116	-0.034410		
27	-0.005944	72	0.002233	117	-0.008176		
28	0.003115	73	-0.012101	118	0.021157		
29	-0.022735	74	0.010285	119	0.029336		
30	-0.007763	75	-0.015259	120	0.011900		
31	-0.021139	76	-0.005800	121	0.026019		
32	-0.015381	77	-0.015313	122	0.009742		
33	0.010643	78	-0.026996	123	-0.009941		
34	-0.002952	79	-0.006391	124	-0.014732		
35	0.032027	80	-0.007820	125	-0.020730		
36	-0.010370	81	-0.024914	126	-0.015615		
37	0.021556	82	-0.003401	127	0.034214		
38	0.004952	83	-0.004807	128	0.014207		
39	0.025840	84	0.002939	129	0.015578		
40	-0.004653	85	-0.005878	130	0.030992		
41	0.023617	86	-0.029838	131	0.000554		
42	0.006880	87	0.007157	132	0.010837		
43	-0.016744	88	0.015125	133	-0.057274		
44	-0.002198	89	-0.001778	134	0.069765		

Table F-30
Estimated Effects for the Mathematics Main Principal Components
Grade 12, Data Analysis, Statistics, and Probability

Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect
0	0.030571	45	0.023903	90	0.055065	135	0.011393
1	0.167328	46	0.031649	91	0.000538	136	-0.015210
2	0.011734	47	-0.003926	92	-0.018682	137	-0.076000
3	0.001109	48	0.013979	93	0.025485	138	-0.061539
4	-0.047681	49	0.014677	94	0.013364		
5	-0.062298	50	-0.048865	95	-0.013614		
6	0.093575	51	-0.011089	96	-0.004227		
7	0.021038	52	-0.000526	97	-0.003508		
8	-0.057759	53	-0.013439	98	0.016964		
9	0.016497	54	-0.020592	99	0.017286		
10	-0.014207	55	0.004162	100	-0.044347		
11	-0.045220	56	0.018788	101	0.014491		
12	-0.086772	57	-0.005194	102	-0.015488		
13	-0.000178	58	0.011849	103	0.057085		
14	0.044614	59	-0.011266	104	-0.033037		
15	-0.003339	60	-0.009519	105	0.017043		
16	-0.051754	61	-0.014162	106	-0.005944		
17	-0.006694	62	0.015297	107	-0.002114		
18	-0.032460	63	0.002032	108	-0.016144		
19	-0.044687	64	0.000176	109	0.010970		
20	0.023884	65	0.020923	110	0.008733		
21	-0.009625	66	-0.015961	111	0.000289		
22	0.007846	67	0.030503	112	-0.013053		
23	0.055882	68	0.007289	113	0.013094		
24	-0.008640	69	-0.005080	114	-0.021386		
25	-0.012036	70	-0.019278	115	0.012062		
26	0.003350	71	-0.018882	116	-0.031140		
27	-0.013707	72	0.015881	117	0.005199		
28	-0.007887	73	-0.022671	118	-0.015140		
29	-0.027128	74	0.017983	119	0.014446		
30	-0.001595	75	0.006524	120	0.007653		
31	-0.017112	76	0.001385	121	0.034102		
32	-0.014832	77	-0.025266	122	0.000678		
33	-0.004783	78	-0.033011	123	0.012045		
34	0.016186	79	-0.007385	124	-0.032659		
35	0.031824	80	-0.000992	125	-0.043477		
36	-0.005076	81	-0.006239	126	-0.019787		
37	0.014360	82	-0.025353	127	0.027614		
38	0.025820	83	0.004464	128	-0.002010		
39	0.027731	84	0.001038	129	-0.015263		
40	-0.032915	85	-0.000691	130	-0.000969		
41	0.026923	86	-0.044789	131	-0.006487		
42	-0.007132	87	-0.009633	132	-0.005609		
43	-0.022201	88	0.014072	133	-0.022876		
44	-0.019310	89	0.003482	134	0.021043		

Table F-31
Estimated Effects for the Mathematics Main Principal Components
Grade 12, Algebra and Functions

Principal Component	Estimated Effect
0	0.020721
1	0.182867
2	0.011940
3	0.004173
4	-0.099940
5	-0.082997
6	0.073254
7	0.051769
8	-0.060409
9	-0.001847
10	-0.018988
11	-0.028550
12	-0.086110
13	0.006875
14	0.028006
15	0.003984
16	-0.015608
17	0.017161
18	0.007373
19	-0.021002
20	0.002782
21	-0.023754
22	0.013023
23	0.017180
24	-0.012560
25	0.021292
26	0.002667
27	0.007417
28	0.004687
29	-0.020093
30	0.008165
31	-0.015377
32	-0.009550
33	0.006251
34	0.010661
35	0.030285
36	-0.006421
37	0.021627
38	0.008231
39	0.031185
40	0.003376
41	0.013028
42	-0.008297
43	-0.016984
44	-0.007041
45	0.006074
46	0.019569
47	0.009725
48	-0.000700
49	0.005942
50	-0.018136
51	-0.032274
52	0.000015
53	-0.017001
54	-0.022970
55	-0.014876
56	0.011830
57	0.004644
58	-0.016472
59	-0.024616
60	-0.004904
61	-0.008428
62	0.016220
63	-0.000362
64	0.005189
65	0.015214
66	0.007814
67	0.013110
68	0.000455
69	0.007332
70	-0.006913
71	0.004794
72	0.013644
73	-0.017615
74	0.004924
75	-0.005929
76	-0.008081
77	-0.000846
78	-0.013812
79	-0.005375
80	0.016239
81	-0.002879
82	-0.021714
83	0.003004
84	-0.002787
85	-0.005053
86	-0.024052
87	0.001580
88	0.006183
89	0.000863
90	0.030031
91	-0.023072
92	-0.029680
93	0.019219
94	0.005797
95	-0.018929
96	-0.015490
97	0.002221
98	0.026666
99	0.001230
100	-0.017615
101	0.005526
102	-0.034616
103	0.021546
104	-0.038941
105	-0.004377
106	0.002526
107	0.014021
108	-0.017691
109	-0.004750
110	0.015362
111	0.002974
112	-0.003002
113	0.002709
114	-0.001006
115	-0.008611
116	-0.040437
117	-0.017891
118	0.013363
119	0.006691
120	0.001246
121	0.010774
122	0.009760
123	-0.005351
124	-0.008766
125	-0.038246
126	0.008895
127	0.007086
128	0.022077
129	-0.008214
130	0.012212
131	-0.006402
132	-0.015493
133	-0.031732
134	0.030811
135	0.016322
136	-0.027487
137	-0.028579
138	0.004592

Table F-32
Estimated Effects for the Writing Main Principal Components
Grade 4, 25-minute Blocks

Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect
0	-0.043209	45	0.012374	90	0.041681
1	0.007596	46	0.001179	91	-0.012722
2	-0.155345	47	-0.026324	92	-0.052127
3	-0.136550	48	0.023466	93	-0.003189
4	0.139432	49	-0.004734	94	0.032953
5	-0.071847	50	-0.033038	95	-0.006753
6	0.018614	51	0.003409	96	0.010507
7	-0.073707	52	-0.031145	97	-0.004486
8	-0.039812	53	0.011475	98	0.009068
9	0.085724	54	0.003051	99	0.033692
10	-0.052640	55	0.016604	100	-0.032182
11	-0.039899	56	-0.011542	101	0.045696
12	0.022905	57	-0.036923	102	0.046162
13	0.029889	58	0.009924	103	0.037590
14	0.080879	59	-0.019675	104	-0.003525
15	-0.017635	60	-0.004745		
16	-0.077088	61	-0.025926		
17	0.003099	62	-0.045278		
18	0.058088	63	-0.038682		
19	-0.020402	64	-0.016530		
20	0.011230	65	-0.024284		
21	-0.003021	66	-0.005379		
22	0.030138	67	0.010572		
23	-0.030652	68	0.003342		
24	-0.017331	69	0.038533		
25	0.048344	70	-0.025561		
26	-0.015887	71	-0.031875		
27	-0.025726	72	-0.015853		
28	0.029784	73	0.012731		
29	0.018436	74	0.040927		
30	0.026760	75	-0.001169		
31	0.033740	76	0.026440		
32	0.003776	77	0.046453		
33	0.034766	78	-0.020512		
34	0.013281	79	0.027976		
35	0.005368	80	-0.029921		
36	0.049354	81	-0.008892		
37	0.004834	82	-0.098857		
38	-0.030752	83	-0.024079		
39	0.046810	84	0.004917		
40	0.012019	85	0.052142		
41	0.022545	86	0.012708		
42	0.010274	87	-0.024349		
43	0.029157	88	-0.034260		
44	0.025742	89	-0.009275		

Table F-33
Estimated Effects for the Writing Main Principal Components
Grade 8, 25-minute Blocks

Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect
0	-0.021987	45	-0.006508	90	0.000819	135	0.022104
1	-0.022591	46	0.002543	91	-0.017206	136	0.041162
2	-0.006368	47	0.007085	92	0.018264	137	0.027865
3	-0.077296	48	0.005105	93	-0.033102	138	-0.003106
4	0.066360	49	0.006255	94	0.017355	139	0.027821
5	0.117349	50	0.003280	95	0.022623	140	-0.038830
6	0.048348	51	0.020189	96	0.010429	141	0.006059
7	-0.031289	52	-0.018560	97	-0.021592	142	0.040644
8	-0.072790	53	-0.001132	98	-0.004989	143	-0.005653
9	0.009052	54	0.036197	99	-0.020950	144	0.026029
10	0.050011	55	0.039040	100	-0.009271	145	0.037750
11	0.107426	56	-0.014019	101	0.000039	146	-0.003699
12	-0.068115	57	-0.024207	102	-0.031197	147	-0.018822
13	0.044045	58	0.000714	103	0.005833	148	0.025347
14	-0.030558	59	0.009022	104	-0.012797	149	0.012907
15	0.030850	60	-0.033581	105	-0.004921	150	0.006418
16	0.004985	61	0.060748	106	0.007306	151	0.046377
17	-0.004716	62	0.005960	107	-0.006375	152	0.024888
18	-0.000086	63	0.017417	108	0.011904	153	0.037729
19	0.001331	64	0.005973	109	-0.006666	154	0.005757
20	-0.009453	65	0.003892	110	0.017707	155	-0.004189
21	-0.002549	66	0.009815	111	0.055879	156	-0.044540
22	0.020373	67	-0.014915	112	-0.001581	156	-0.044540
23	0.023661	68	-0.027769	113	0.003192	157	0.038451
24	-0.024640	69	0.004865	114	-0.019801	158	0.022430
25	-0.043578	70	0.004276	115	-0.013192	159	-0.008646
26	0.032245	71	-0.018657	116	-0.020487	160	-0.028468
27	-0.012443	72	-0.012737	117	0.011448	161	0.022723
28	-0.054345	73	0.018621	118	-0.024088	162	-0.026912
29	0.100348	74	-0.010455	119	-0.008358	163	0.027588
30	0.008090	75	-0.002489	120	-0.015886	164	0.000774
31	-0.021028	76	0.002598	121	-0.003088	165	0.015943
32	-0.030649	77	0.006244	122	0.011096	166	-0.026013
33	0.017014	78	-0.001847	123	0.007583	167	-0.014637
34	0.021680	79	-0.036374	124	0.037643	168	-0.009844
35	0.054761	80	-0.001068	125	0.005199	169	0.085152
36	-0.002611	81	0.019508	126	-0.026095	170	-0.101697
37	-0.000371	82	-0.038817	127	0.009371	171	-0.075634
38	0.010406	83	-0.010985	128	0.003090	172	-0.038610
39	-0.012698	84	0.016043	129	0.003311	173	-0.026957
40	0.029844	85	0.024475	130	0.005457	174	-0.008120
41	-0.027857	86	0.024133	131	0.007577	175	0.023556
42	0.003853	87	-0.015944	132	0.019722	176	-0.014161
43	-0.053129	88	-0.017223	133	0.005961	177	0.057034
44	0.000761	89	0.027326	134	0.006353	178	-0.019436

Table F-33 (continued)
Estimated Effects for the Writing Main Principal Components
Grade 8, 25-minute Blocks

Principal Component	Estimated Effect
179	0.006563
180	0.028802
181	0.018332
182	-0.000365
183	0.028078
184	0.072809
185	-0.003385
186	-0.034837
187	0.006621
188	0.029645
189	-0.004694
190	0.024015
191	-0.014413
192	0.031355
193	0.009086
194	0.040220
195	-0.015122
196	-0.022874
197	0.006271
198	-0.004629
199	-0.028414
200	-0.011612
201	-0.041798
202	0.008457
203	-0.035334
204	-0.030423
205	-0.001563
206	0.011142

Table F-34
Estimated Effects for the Writing Main Principal Components
Grade 8, 50-minute Blocks

Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect
0	-0.036677	45	-0.004577	90	0.011952	135	0.031659
1	-0.025012	46	-0.018292	91	-0.023229	136	0.052007
2	-0.001730	47	-0.016629	92	0.036816	137	0.015996
3	-0.062578	48	0.015979	93	-0.018648	138	-0.004464
4	0.057945	49	0.012422	94	-0.026399	139	0.001871
5	0.124370	50	-0.021390	95	0.011091	140	-0.068821
6	0.070661	51	0.031065	96	0.041979	141	-0.038965
7	-0.050414	52	-0.040259	97	-0.018946	142	0.007421
8	-0.100537	53	-0.020535	98	0.013652	143	-0.049990
9	0.022875	54	0.034747	99	-0.009635	144	0.067382
10	0.041465	55	0.031057	100	-0.036803	145	0.040997
11	0.093182	56	-0.037402	101	-0.001215	146	0.018518
12	-0.039384	57	0.012179	102	-0.023927	147	-0.004248
13	0.016409	58	0.003345	103	0.021817	148	0.048071
14	-0.039254	59	0.036921	104	0.016563	149	0.012923
15	0.025434	60	-0.018044	105	0.019738	150	0.042774
16	-0.034164	61	-0.008650	106	0.037250	151	0.021412
17	0.018601	62	0.028698	107	-0.009259	152	0.024851
18	-0.020368	63	-0.006531	108	0.035684	153	0.031927
19	-0.003009	64	0.002434	109	0.004371	154	-0.011696
20	-0.010318	65	0.013145	110	-0.005553	155	-0.062106
21	0.004810	66	0.025222	111	0.060933	156	-0.037299
22	0.033424	67	-0.014303	112	0.022835	157	0.011542
23	0.011309	68	-0.032359	113	0.010193	158	0.058643
24	-0.048144	69	-0.027321	114	0.048783	159	-0.022581
25	-0.041923	70	-0.022049	115	0.000652	160	0.027675
26	0.002600	71	-0.028014	116	-0.004565	161	-0.018062
27	0.024765	72	0.004893	117	-0.005924	162	-0.013594
28	-0.043833	73	0.017286	118	-0.036513	163	0.029810
29	0.119564	74	-0.015190	119	-0.021528	164	-0.011976
30	0.024988	75	-0.031464	120	-0.011826	165	0.049975
31	-0.014017	76	0.004535	121	-0.025633	166	-0.015294
32	-0.020417	77	0.001397	122	-0.007010	167	-0.022994
33	0.019021	78	0.002572	123	-0.005371	168	-0.033649
34	0.001702	79	0.004758	124	-0.008048	169	0.012935
35	0.049044	80	0.006372	125	0.006758	170	-0.031282
36	-0.008201	81	0.055591	126	0.000805	171	-0.028945
37	0.011978	82	0.001869	127	-0.010066	172	-0.073630
38	-0.024582	83	-0.007400	128	0.023538	173	-0.063847
39	0.001261	84	0.022092	129	0.004489	174	-0.021669
40	-0.006305	85	-0.016893	130	0.039119	175	0.030920
41	-0.056755	86	0.059181	131	-0.010821	176	0.054561
42	0.014033	87	-0.031103	132	-0.030989	177	0.064613
43	-0.067474	88	-0.009043	133	0.022070	178	-0.043962
44	0.039972	89	0.037539	134	-0.043961	179	0.022466

Table F-34 (continued)
Estimated Effects for the Writing Main Principal Components
Grade 8, 50-minute Blocks

Principal Component	Estimated Effect
180	0.014722
181	0.035793
182	-0.000639
183	-0.016471
184	0.025033
185	-0.037981
186	-0.020397
187	0.027283
188	-0.002130
189	0.003991
190	-0.010498
191	-0.040302
192	0.016923
193	-0.048757
194	0.094776
195	0.002020
196	0.045924
197	-0.019775
198	0.004134
199	0.011048
200	-0.004199
201	0.033428
202	0.031256
203	-0.045268
204	-0.012336
205	0.001025
206	-0.010148

Table F-35
Estimated Effects for the Writing Main Principal Components
Grade 12, 25-minute Blocks

Principal Component	Estimated Effect	Principal Component	Estimated Effect	Principal Component	Estimated Effect
0	-0.025140	45	-0.045201	90	0.010800
1	-0.056620	46	-0.002450	91	0.036635
2	0.016738	47	0.005716	92	0.000297
3	0.158670	48	0.005719	93	0.026583
4	0.158624	49	0.006081	94	-0.007620
5	-0.082124	50	-0.040745	95	0.004963
6	-0.021727	51	-0.014002	96	0.014474
7	0.104365	52	0.019455	97	-0.006753
8	0.015057	53	-0.026328	98	-0.035687
9	-0.009896	54	-0.005996	99	-0.000948
10	-0.051603	55	-0.022279	100	-0.024081
11	0.103924	56	0.012136	101	0.010273
12	0.017101	57	0.028207	102	0.005376
13	-0.044107	58	0.024848	103	0.007870
14	0.032681	59	0.027159	104	-0.057762
15	0.005730	60	0.004503	105	-0.025573
16	0.028834	61	-0.023481	106	0.001002
17	-0.074515	62	0.025907	107	-0.004004
18	-0.012128	63	-0.008974	108	-0.008971
19	0.009252	64	-0.024949	109	0.007074
20	-0.014358	65	-0.007819	110	0.030700
21	-0.006669	66	-0.019970	111	0.015518
22	0.001387	67	-0.013585	112	-0.004865
23	-0.003915	68	0.005130	113	-0.009226
24	-0.027030	69	0.016261	114	0.103977
25	-0.000860	70	0.011825	115	-0.004605
26	-0.034674	71	-0.012276	116	0.058947
27	-0.000490	72	0.021269	117	0.032332
28	-0.055091	73	-0.006448	118	-0.041697
29	-0.014761	74	0.018238	119	-0.003994
30	-0.013217	75	-0.038455	120	-0.001860
31	0.026541	76	0.000762		
32	-0.055906	77	0.003197		
33	0.021389	78	-0.015773		
34	-0.029163	79	0.039154		
35	0.011547	80	-0.006355		
36	0.024485	81	0.027227		
37	0.003025	82	-0.025045		
38	0.036685	83	-0.006668		
39	0.036089	84	-0.017597		
40	-0.018365	85	-0.027548		
41	-0.022138	86	-0.041433		
42	-0.044324	87	-0.056968		
43	-0.010834	88	0.019246		
44	-0.000497	89	0.011906		

Table F-36
Estimated Effects for the Writing Main Principal Components
Grade 12, 50-minute Blocks

Principal Component	Estimated Effect
0	-0.010192
1	-0.050984
2	0.006041
3	0.169433
4	0.138099
5	-0.082838
6	-0.008294
7	0.120801
8	0.028851
9	-0.023161
10	-0.064909
11	0.118666
12	0.049106
13	-0.072703
14	0.024687
15	0.028115
16	0.051833
17	-0.060840
18	-0.029154
19	0.007575
20	0.017422
21	0.012857
22	-0.000787
23	-0.001166
24	-0.046190
25	-0.009189
26	-0.029291
27	0.000506
28	-0.018337
29	0.035548
30	0.001219
31	0.020220
32	-0.048180
33	0.004098
34	0.006302
35	0.019698
36	0.004856
37	0.000579
38	0.018470
39	0.038827
40	-0.008922
41	-0.039625
42	-0.000020
43	-0.012190
44	-0.004906
45	-0.001539
46	0.003035
47	0.029471
48	-0.001713
49	-0.002320
50	-0.036195
51	0.021864
52	0.000211
53	-0.008470
54	-0.002775
55	-0.011504
56	0.000660
57	0.014853
58	-0.008941
59	0.034486
60	0.002656
61	-0.014713
62	-0.003940
63	0.008236
64	-0.022135
65	-0.019551
66	-0.016058
67	0.023473
68	-0.004471
69	0.018862
70	-0.031049
71	0.001028
72	0.016070
73	-0.011857
74	0.009868
75	-0.018969
76	0.021392
77	0.022324
78	0.034239
79	0.021633
80	0.009524
81	0.031799
82	0.009244
83	-0.028590
84	-0.013765
85	-0.014958
86	-0.025110
87	-0.038599
88	-0.005364
89	0.021447
90	0.023046
91	-0.004057
92	-0.029163
93	0.034192
94	0.010325
95	0.000374
96	-0.018401
97	-0.026794
98	-0.015533
99	-0.008621
100	-0.004519
101	0.015766
102	0.017673
103	0.044312
104	-0.009824
105	-0.024527
106	-0.014151
107	0.001994
108	-0.041434
109	-0.011125
110	0.026312
111	0.071643
112	-0.036842
113	-0.020247
114	0.082832
115	0.019970
116	0.033047
117	0.048756
118	-0.027512
119	-0.011340
120	0.009721

Table F-37
Estimated Effects for Reading Trend Conditioning Variable Contrasts, Age 9

Contrast	Estimated Effect	Description
OVERALL	-0.513732	OVERALL CONSTANT
GENDER2	0.181526	SEX (FEMALE)
ETHNIC2	-0.405613	OBSERVED ETHNICITY (BLACK)
ETHNIC3	-0.293605	OBSERVED ETHNICITY (HISPANIC)
ETHNIC4	0.153435	OBSERVED ETHNICITY (ASIAN)
STOC2	0.304684	SIZE AND TYPE OF COMMUNITY (HIGH METRO)
STOC3	0.130213	SIZE AND TYPE OF COMMUNITY (NOT HIGH OR LOW)
REGION2	-0.185309	REGION (SOUTHEAST)
REGION3	-0.102441	REGION (CENTRAL)
REGION4	-0.135029	REGION (WEST)
PARED2	0.172488	PARENTS EDUCATION (HIGH SCHOOL GRAD)
PARED3	0.186263	PARENTS EDUCATION (POST HIGH SCHOOL)
PARED4	0.318926	PARENTS EDUCATION (COLLEGE GRAD)
PARED_	0.072067	PARENTS EDUCATION (MISSING, I DON'T KNOW)
TV. L	-0.064756	TV WATCHING (LINEAR)
TV. _	-0.373340	TV WATCHING (MISSING)
HOMEITML	0.089096	ARTICLES IN HOME (6) (LINEAR 3)
HOMEITM_	-0.668447	ARTICLES IN HOME (6) (MISSING)
HW-LIN	-0.067036	HOMEWORK (LINEAR)
HW-???	-0.608807	HOMEWORK (MISSING)
HOMELNG1	-0.263637	OTHER LANGUAGE AT HOME (SPANISH/OTHER)
HOMELNG2	-0.231772	OTHER LANGUAGE AT HOME (MISSING)
PGSREAD1	0.129474	PAGES READ (6 OR MORE)
PGSREAD_	-0.110396	PAGES READ (MISSING)
%WHITE 1	0.183571	PERCENT WHITE 0-49%
%WHITE 2	0.184486	PERCENT WHITE 50-79%
%WHITE 3	0.239537	PERCENT WHITE 80-100%
AGE/GRD1	-0.425889	MODAL AGE, < MODAL GRADE
AGE/GRD2	0.220819	MODAL AGE, MODAL GRADE, MISSING
AGE/GRD3	0.796383	MODAL AGE, > MODAL GRADE
AGE/GRD4	0.050720	> MODAL AGE, MODAL GRADE
%LUNCH 2	0.073338	PERCENT IN LUNCH PROGRAM 2
%LUNCH 3	0.035287	PERCENT IN LUNCH PROGRAM 3
%LUNCH 4	-0.057456	PERCENT IN LUNCH PROGRAM 4
%LUNCH 5	-0.108688	PERCENT IN LUNCH PROGRAM 5
%LUNCH 6	-0.223507	PERCENT IN LUNCH PROGRAM 6
%LUNCH 7	-0.331788	PERCENT IN LUNCH PROGRAM 7
%LUNCH 8	-0.293972	PERCENT IN LUNCH PROGRAM 8
%LUNCH 9	-0.160247	MISSING LUNCH PROGRAM
CRSTAKNL	0.008699	COURSES TAKEN LINEAR (N OF 7)
CRSTAKNM	-0.159933	COURSES TAKEN MISSING

Table F-38
Estimated Effects for Reading Trend Conditioning Variable Contrasts, Age 13

Contrast	Estimated Effect	Description
OVERALL	-0.220259	OVERALL CONSTANT
GENDER2	0.174846	SEX (FEMALE)
ETHNIC2	-0.397093	OBSERVED ETHNICITY (BLACK)
ETHNIC3	-0.199209	OBSERVED ETHNICITY (HISPANIC)
ETHNIC4	0.246872	OBSERVED ETHNICITY (ASIAN)
STOC2	0.255913	SIZE AND TYPE OF COMMUNITY (HIGH METRO)
STOC3	0.184486	SIZE AND TYPE OF COMMUNITY (NOT HIGH OR LOW)
REGION2	0.005833	REGION (SOUTHEAST)
REGION3	0.105127	REGION (CENTRAL)
REGION4	-0.002101	REGION (WEST)
PARED2	0.114411	PARENTS EDUCATION (HIGH SCHOOL GRAD)
PARED3	0.318317	PARENTS EDUCATION (POST HIGH SCHOOL)
PARED4	0.398001	PARENTS EDUCATION (COLLEGE GRAD)
PARED_	-0.046357	PARENTS EDUCATION (MISSING, I DON'T KNOW)
TV. L	-0.033673	TV WATCHING (LINEAR)
TV. _	-0.445422	TV WATCHING (Missing)
HOMEITML	0.093448	Articles IN HOME (6) (Linear 3)
HOMEITM_	0.688006	ARTICLES IN HOME (6) (Missing)
HW-Lin	0.053756	HOMEWORK (Linear)
HW-???	-0.427362	HOMEWORK (Missing)
HOMELNG1	-0.192932	OTHER LANGUAGE AT HOME (Spanish/Other)
HOMELNG2	-0.369118	OTHER LANGUAGE AT HOME (Missing)
PGSREAD1	0.179001	PAGES READ (6 OR MORE)
PGSREAD_	-0.021963	PAGES READ (MISING)
%WHITE 1	-0.108630	PERCENT WHITE 0-49%
%WHITE 2	-0.134093	PERCENT WHITE 50-79%
%WHITE 3	-0.143479	PERCENT WHITE 80-100%
AGE/GRD1	-0.815620	MODAL AGE, < MODAL GRADE
AGE/GRD2	-0.313020	MODAL AGE, MODAL GRADE, MISSING
AGE/GRD3	0.552151	MODAL AGE, > MODAL GRADE
AGE/GRD4	-0.572083	> MODAL AGE, MODAL GRADE
%LUNCH 2	0.003414	PERCENT IN LUNCH PROGRAM 2
%LUNCH 3	-0.061430	PERCENT IN LUNCH PROGRAM 3
%LUNCH 4	-0.147146	PERCENT IN LUNCH PROGRAM 4
%LUNCH 5	-0.161726	PERCENT IN LUNCH PROGRAM 5
%LUNCH 6	-0.232102	PERCENT IN LUNCH PROGRAM 6
%LUNCH 7	-0.218322	PERCENT IN LUNCH PROGRAM 7
%LUNCH 8	-0.256407	PERCENT IN LUNCH PROGRAM 8
%LUNCH 9	0.157709	MISSING LUNCH PROGRAM
CRSTAKNL	0.033123	COURSES TAKEN LINEAR (n of 7)
CRSTAKNM	0.001323	COURSES TAKEN MISSING

Table F-39
Estimated Effects for Reading Trend Conditioning Variable Contrasts, Age 17

Contrast	Estimated Effect	Description
OVERALL	-0.340466	OVERALL CONSTANT
GENDER2	0.156699	SEX (FEMALE)
ETHNIC2	-0.484438	OBSERVED ETHNICITY (BLACK)
ETHNIC3	-0.269971	OBSERVED ETHNICITY (HISPANIC)
ETHNIC4	-0.063550	OBSERVED ETHNICITY (ASIAN)
STOC2	0.248231	SIZE AND TYPE OF COMMUNITY (HIGH METRO)
STOC3	0.088602	SIZE AND TYPE OF COMMUNITY (NOT HIGH OR LOW)
REGION2	-0.078506	REGION (SOUTHEAST)
REGION3	-0.087270	REGION (CENTRAL)
REGION4	-0.022823	REGION (WEST)
PARED2	-0.087612	PARENTS EDUCATION (HIGH SCHOOL GRAD)
PARED3	0.094277	PARENTS EDUCATION (POST HIGH SCHOOL)
PARED4	0.173957	PARENTS EDUCATION (COLLEGE GRAD)
PARED_	-0.248750	PARENTS EDUCATION (MISSING, I DON'T KNOW)
TV. L	-0.082505	TV WATCHING (LINEAR)
TV. _	-1.032357	TV WATCHING (MISSING)
HOMEITML	0.114223	ARTICLES IN HOME (6) (LINEAR 3)
HOMEITM_	0.815660	ARTICLES IN HOME (6) (MISSING)
HW-LIN	0.088894	HOMEWORK (LINEAR)
HW-???	-0.683952	HOMEWORK (MISSING)
HOMELNG1	-0.260019	OTHER LANGUAGE AT HOME (SPANISH/OTHER)
HOMELNG2	-0.523342	OTHER LANGUAGE AT HOME (MISSING)
PGSREAD1	0.295160	PAGES READ (6 OR MORE)
PGSREAD_	-0.130337	PAGES READ (MISSING)
%WHITE 1	0.162065	PERCENT WHITE 0-49%
%WHITE 2	0.034887	PERCENT WHITE 50-79%
%WHITE 3	0.009797	PERCENT WHITE 80-100%
AGE/GRD1	-0.656412	MODAL AGE, < MODAL GRADE
AGE/GRD2	-0.064569	MODAL AGE, MODAL GRADE, MISSING
AGE/GRD3	-0.009247	MODAL AGE, > MODAL GRADE
AGE/GRD4	-0.442310	> MODAL AGE, MODAL GRADE
%LUNCH 2	-0.162961	PERCENT IN LUNCH PROGRAM 2
%LUNCH 3	-0.097485	PERCENT IN LUNCH PROGRAM 3
%LUNCH 4	-0.216100	PERCENT IN LUNCH PROGRAM 4
%LUNCH 5	-0.278840	PERCENT IN LUNCH PROGRAM 5
%LUNCH 6	-0.219360	PERCENT IN LUNCH PROGRAM 6
%LUNCH 7	-0.177751	PERCENT IN LUNCH PROGRAM 7
%LUNCH 8	-0.679125	PERCENT IN LUNCH PROGRAM 8
%LUNCH 9	-0.147992	MISSING LUNCH PROGRAM
CRSTAKNL	0.000000	COURSES TAKEN LINEAR (N OF 7)
CRSTAKNM	0.000000	COURSES TAKEN MISSING

Table F-40
Estimated Effects for Mathematics Trend Conditioning Variable Contrasts, Age 9

Contrast	Estimated Effect	Description
OVERALL	0.197876	OVERALL CONSTANT
GENDER2	-0.117454	SEX (FEMALE)
ETHNIC2	-0.176928	OBSERVED ETHNICITY (BLACK)
ETHNIC3	-0.214051	OBSERVED ETHNICITY (HISPANIC)
ETHNIC4	0.380647	OBSERVED ETHNICITY (ASIAN)
STOC3	-0.198079	SIZE AND TYPE OF COMMUNITY (HIGH METRO)
STOC1	0.270845	SIZE & TYPE OF COMM (NOT HI, NOT LOW)
REGION2	-0.095139	REGION (SOUTHEAST)
REGION3	0.018361	REGION (CENTRAL)
REGION4	-0.048945	REGION (WEST)
PARED2	0.069762	PARENTS EDUCATION (HIGH SCHOOL GRAD)
PARED3	0.336209	PARENTS EDUCATION (POST HIGH SCHOOL)
PARED4	0.284043	PARENTS EDUCATION (COLLEGE GRAD)
PARED_	0.064992	PARENTS EDUCATION (MISSING,I DON'T KNOW)
MODGRD 1	-0.914331	< MODAL GRADE
MODGRD 2	0.480935	> MODAL GRADE
HOMEEN22	0.204375	3 ITEMS IN THE HOME
HOMEEN23	0.343383	4 ITEMS IN THE HOME
RAC/SEX6	-0.105173	BLACK, FEMALE
RAC/SEX7	0.132975	HISPANIC, FEMALE
RAC/SEX8	0.163511	ASIAN AMERICAN, FEMALE
RAC/PED7	-0.145025	BLACK, HIGH SCHOOL GRAD/POST HS
RAC/PED	0.000000	BLACK, POST HIGH SCHOOL
RAC/PED8	-0.297880	BLACK, COLLEGE GRAD
RAC/PED9	-0.060233	BLACK, MISSING
RAC/PE11	-0.131073	HISPANIC, HIGH SCHOOL GRAD/POST HS
RAC/PED	0.000000	HISPANIC, POST HIGH SCHOOL
RAC/PE12	-0.173323	HISPANIC, COLLEGE GRAD
RAC/PE13	-0.158954	HISPANIC, MISSING
RAC/PE15	-0.950083	ASIAN AMERICAN, HIGH SCH GRAD/POST HS
RAC/PED	0.000000	ASIAN AMERICAN, POST HS
RAC/PE16	-0.543517	ASIAN AMERICAN, COLLEGE GRAD
RAC/PE17	-0.696027	ASIAN AMERICAN, MISSING
SCHTYPE	0.070063	SCHOOL (NONPUBLIC)
LANGHOM1	0.006899	OTHER THAN ENGLISH AT HOME (SOMETIMES)
LANGHOM2	-0.306969	OTHER THAN ENGLISH AT HOME (ALWAYS)
RAC/LNG1	-0.027845	BLACK, ALWAYS OTHER THAN ENGLISH
RAC/LNG2	0.000000	BLACK, SOMETIMES OTHER THAN ENGLISH
RAC/LNG3	0.197098	HISPANIC, ALWAYS OTHER THAN ENGLISH
RAC/LNG4	0.000000	HISPANIC, SOMETIMES OTHER THAN ENGLISH
RAC/LNG5	0.612691	ASIAN AM, ALWAYS OTHER THAN ENGLISH
RAC/LNG6	0.000000	ASIAN AM, SOMETIMES OTHER THAN ENGLISH
DRACE2	-0.137113	DERIVED RACE (BLACK)
DRACE3	-0.328998	DERIVED RACE (HISPANIC)
DRACE4	-0.104348	DERIVED RACE (ASIAN AMERICAN)

Table F-41
Estimated Effects for Mathematics Trend Conditioning Variable Contrasts, Age 13

Contrast	Estimated Effect	Description
OVERALL	-0.158888	OVERALL CONSTANT
GENDER2	-0.186400	SEX (FEMALE)
ETHNIC2	-0.666451	OBSERVED ETHNICITY (BLACK)
ETHNIC3	0.013604	OBSERVED ETHNICITY (HISPANIC)
ETHNIC4	-0.001169	OBSERVED ETHNICITY (ASIAN)
STOC3	-0.261531	SIZE AND TYPE OF COMMUNITY (HIGH METRO)
STOC1	0.279924	SIZE & TYPE OF COMM (NOT HI, NOT LOW)
REGION2	0.119508	REGION (SOUTHEAST)
REGION3	0.097647	REGION (CENTRAL)
REGION4	-0.001382	REGION (WEST)
PARED2	0.076253	PARENTS EDUCATION (HIGH SCHOOL GRAD)
PARED3	0.371673	PARENTS EDUCATION (POST HIGH SCHOOL)
PARED4	0.516345	PARENTS EDUCATION (COLLEGE GRAD)
PARED_	-0.133604	PARENTS EDUCATION (MISSING,I DON'T KNOW)
MODGRD 1	-0.588800	< MODAL GRADE
MODGRD 2	1.086030	> MODAL GRADE
HOMEEN22	0.150945	3 ITEMS IN THE HOME
HOMEEN23	0.264431	4 ITEMS IN THE HOME
RAC/SEX6	0.100425	BLACK, FEMALE
RAC/SEX7	0.066662	HISPANIC, FEMALE
RAC/SEX8	0.146834	ASIAN AMERICAN, FEMALE
RAC/PED7	0.035197	BLACK, HIGH SCHOOL GRAD
RAC/PED8	-0.042410	BLACK, POST HIGH SCHOOL
RAC/PED9	-0.211362	BLACK, COLLEGE GRAD
RAC/PE10	-0.037304	BLACK, MISSING
RAC/PE12	-0.110075	HISPANIC, HIGH SCHOOL GRAD
RAC/PE13	0.132611	HISPANIC, POST HIGH SCHOOL
RAC/PE14	-0.212220	HISPANIC, COLLEGE GRAD
RAC/PE15	0.009007	HISPANIC, MISSING
RAC/PE17	0.192939	ASIAN AMERICAN, HIGH SCH GRAD
RAC/PE18	-0.727266	ASIAN AMERICAN, POST HIGH SCHOOL
RAC/PE19	-0.418656	ASIAN AMERICAN, COLLEGE GRAD
RAC/PE20	-0.180341	ASIAN AMERICAN, MISSING
SCHTYPE	0.017616	SCHOOL (NONPUBLIC)
LANGHOM1	0.129548	OTHER THAN ENGLISH AT HOME (SOMETIMES)
LANGHOM2	-0.189226	OTHER THAN ENGLISH AT HOME (ALWAYS)
RAC/LNG1	-0.064344	BLACK, ALWAYS OTHER THAN ENGLISH
RAC/LNG2	0.007464	BLACK, SOMETIMES OTHER THAN ENGLISH
RAC/LNG3	0.200808	HISPANIC, ALWAYS OTHER THAN ENGLISH
RAC/LNG4	-0.050840	HISPANIC, SOMETIMES OTHER THAN ENGLISH
RAC/LNG5	0.536255	ASIAN AM, ALWAYS OTHER THAN ENGLISH
RAC/LNG6	0.438875	ASIAN AM, SOMETIMES OTHER THAN ENGLISH
DRACE2	0.071386	DERIVED RACE (BLACK)
DRACE3	-0.408811	DERIVED RACE (HISPANIC)

Table F-41 (continued)
Estimated Effects for Mathematics Trend Conditioning Variable Contrasts, Age 13

Contrast	Estimated Effect	Description
DRACE4	0.187747	DERIVED RACE (ASIAN AMERICAN)
HW-NO	-0.050043	HOMEWORK (NONE ASSIGNED)
HW-YES	-0.115781	HOMEWORK (YES)
HW-234	0.078139	HOMEWORK (1/2 HR TO 2 HOURS)

Table F-42
Estimated Effects for Mathematics Trend Conditioning Variable Contrasts, Age 17

Contrast	Estimated Effect	Description
OVERALL	-0.650121	OVERALL CONSTANT
GENDER2	-0.236919	SEX (FEMALE)
ETHNIC2	-0.604783	OBSERVED ETHNICITY (BLACK)
ETHNIC3	0.066467	OBSERVED ETHNICITY (HISPANIC)
ETHNIC4	-0.043076	OBSERVED ETHNICITY (ASIAN)
STOC3	-0.119068	SIZE AND TYPE OF COMMUNITY (HIGH METRO)
STOC1	0.143214	SIZE & TYPE OF COMM (NOT HI, NOT LOW)
REGION2	0.031454	REGION (SOUTHEAST)
REGION3	0.042038	REGION (CENTRAL)
REGION4	-0.066320	REGION (WEST)
PARED2	0.010052	PARENTS EDUCATION (HIGH SCHOOL GRAD)
PARED3	0.066746	PARENTS EDUCATION (POST HIGH SCHOOL)
PARED4	0.176006	PARENTS EDUCATION (COLLEGE GRAD)
PARED_	-0.055532	PARENTS EDUCATION (MISSING,I DON'T KNOW)
MODGRD 1	-0.225309	< MODAL GRADE
MODGRD 2	-0.026764	> MODAL GRADE
HOMEEN22	0.107279	3 ITEMS IN THE HOME
HOMEEN23	0.076838	4 ITEMS IN THE HOME
RAC/SEX6	0.045128	BLACK, FEMALE
RAC/SEX7	0.014569	HISPANIC, FEMALE
RAC/SEX8	0.177639	ASIAN AMERICAN, FEMALE
RAC/PED7	0.224756	BLACK, HIGH SCHOOL GRAD
RAC/PED8	0.412367	BLACK, POST HIGH SCHOOL
RAC/PED9	-0.014418	BLACK, COLLEGE GRAD
RAC/PE10	0.119644	BLACK, MISSING
RAC/PE12	0.061402	HISPANIC, HIGH SCHOOL GRAD
RAC/PE13	0.118558	HISPANIC, POST HIGH SCHOOL
RAC/PE14	-0.170025	HISPANIC, COLLEGE GRAD
RAC/PE15	-0.177833	HISPANIC, MISSING
RAC/PE17	0.211884	ASIAN AMERICAN, HIGH SCH GRAD
RAC/PE18	0.358422	ASIAN AMERICAN, POST HIGH SCH OR COL GRD
RAC/PED	0.000000	ASIAN AMERICAN, EMPTY--COL GRD COLLAPSED
RAC/PE19	0.405039	ASIAN AMERICAN, MISSING
SCHTYPE	-0.037221	SCHOOL (NONPUBLIC)
LANGHOM1	-0.050179	OTHER THAN ENGLISH AT HOME (SOMETIMES)
LANGHOM2	-0.238617	OTHER THAN ENGLISH AT HOME (ALWAYS)
RAC/LNG1	0.019805	BLACK, ALWAYS OTHER THAN ENGLISH
RAC/LNG2	0.090146	BLACK, SOMETIMES OTHER THAN ENGLISH
RAC/LNG3	0.243673	HISPANIC, ALWAYS OTHER THAN ENGLISH
RAC/LNG4	0.214264	HISPANIC, SOMETIMES OTHER THAN ENGLISH
RAC/LNG5	-0.023288	ASIAN AM, ALWAYS OTHER THAN ENGLISH
RAC/LNG6	-0.138054	ASIAN AM, SOMETIMES OTHER THAN ENGLISH
DRACE2	-0.205615	DERIVED RACE (BLACK)
DRACE3	-0.381056	DERIVED RACE (HISPANIC)

Table F-42 (continued)
Estimated Effects for Mathematics Trend Conditioning Variable Contrasts, Age 17

Contrast	Estimated Effect	Description
DRACE4	-0.038334	DERIVED RACE (ASIAN AMERICAN)
HW-NO	-0.220374	HOMEWORK (NONE ASSIGNED)
HW-YES	-0.098911	HOMEWORK (YES)
HW-234	0.018137	HOMEWORK (1/2 HR TO 2 HOURS)
NMATH1	-0.131180	PRE-ALGEBRA
NMATH2	0.357598	ALGEBRA
NMATH3	0.602370	GEOMETRY
NMATH4	1.036484	ALGEBRA 2
NMATH5	1.669434	CALCULUS
HS_PGM2	0.273281	HS PROG - COLL PREP
HS_PGM3	-0.055120	HS PROG - VOCAT/TECH

Table F-43
Estimated Effects for Writing Trend Conditioning Variable Contrasts, Age 9

Contrast	Estimated Effect	Description
OVERALL	-1.881196	OVERALL
FEMALE	0.379198	GENDER: FEMALE
WHIT/OTH	0.560636	ETHNICITY: EVERYONE EXCL BLK, HISP
HISPANIC	0.376564	ETHNICITY: HISPANIC
HI_METRO	0.442727	STOC: HIGH METROPOLITAN
STOC-OTH	0.046562	STOC: OTHER
S_EAST	-0.302567	REGION: SOUTHEAST
CENTRAL	-0.142851	REGION: CENTRAL
WEST	-0.210188	REGION: WEST
HS_GRAD	0.118596	PARED: HIGH SCHOOL GRADUATE
POST_HS	0.047132	PARED: POST HIGH SCHOOL
COL_GRAD	0.167277	PARED: COLLEGE GRADUATE
PARED-?	0.091873	PARED: MISSING, I DON'T KNOW
GR08/A13	0.000000	GRADE 08 / AGE 13
GR11/A17	0.000000	GRADE 11 / AGE 17
ARTICL=4	0.162959	NUMBER OF HOME ARTICLES : 4
ARTICL=5	0.243481	NUMBER OF HOME ARTICLES : 5
<MODAGE	-0.513983	< MODAL AGE
>MODAGE	-0.079102	> MODAL AGE
HWK:NONE	-0.214862	TIME ON HOMEWORK: NONE DONE
HWK:<1	0.047103	TIME ON HOMEWORK: < 1 HOUR
HWK:1-2	0.056338	TIME ON HOMEWORK: 1-2 HOURS
HWK:>2	-0.166907	TIME ON HOMEWORK: > 2 HOURS
HWK:MISS	-0.111159	TIME ON HOMEWORK: MISSING
NON_PUBL	0.087546	PRIVATE, CATHOLIC, BIA, DOD MISSING
TV-LINER	0.148563	HOURS OF TV WATCH (LINEAR) 0-6
TV-QUAD	-0.028582	HOURS OF IV WATCH (QUAD) 0-36
MOTH_WKS	0.128463	MOTHER WORKS OUTSIDE HOME
LANG_MIN	-0.208438	LANGUAGE MINORITY
LANG_MIS	-0.386669	LANGUAGE MINORITY: MISSING
%LN1-5	0.160918	PERCENT IN LUNCH PROGRAM: 1-5%
%LN6-10	0.030162	PERCENT IN LUNCH PROGRAM: 6-10%
%LN11-15	0.031011	PERCENT IN LUNCH PROGRAM: 11-15%
%LN16-50	-0.052193	PERCENT IN LUNCH PROGRAM: 16-50%
%LN51-75	-0.039754	PERCENT IN LUNCH PROGRAM: 51-75%
%LN76-90	-0.067356	PERCENT IN LUNCH PROGRAM: 76-90%
%LN>90	0.043797	PERCENT IN LUNCH PROGRAM: OVER 90%
INTEGSCH	0.062305	INTEGRATED SCHOOL (50-79.9% WHITE)
WHIT_MIS	0.113741	WHITE (80-100%) OR MISSING
GRINSCHL	0.268436	GRADES IN SCHOOL
11+PAGES	0.290301	PAGES READ: 11 + PAGES
6-10PAGE	0.264520	PAGES READ: 6-10 PAGES
1-5PAGES	0.138874	PAGES READ: 1-5 PAGES
#REPORTS	-0.032271	NUMBER OF ESSAYS, REPORTS PRODUCED
<MODGRD	-0.750254	< MODAL GRADE
>MODGRD	0.716376	> MODAL GRADE

Table F-44
Estimated Effects for Writing Trend Conditioning Variable Contrasts, Age 13

Contrast	Estimated Effect	Description
OVERALL	-1.606148	OVERALL
FEMALE	0.331676	GENDER: FEMALE
WHIT/OTH	0.194883	ETHNICITY: EVERYONE EXCL BLK, HISP
HISPANIC	0.238042	ETHNICITY: HISPANIC
HI_METRO	0.430910	STOC: HIGH METROPOLITAN
STOC-OTH	0.278088	STOC: OTHER
S_EAST	-0.263003	REGION: SOUTHEAST
CENTRAL	-0.062793	REGION: CENTRAL
WEST	-0.236503	REGION: WEST
HS_GRAD	0.101406	PARED: HIGH SCHOOL GRADUATE
POST_HS	0.228343	PARED: POST HIGH SCHOOL
COL_GRAD	0.175858	PARED: COLLEGE GRADUATE
PARED-?	-0.181363	PARED: MISSING, I DON'T KNOW
GR08/A13	0.000000	GRADE 08 / AGE 13
GR11/A17	0.000000	GRADE 11 / AGE 17
ARTICL=4	0.010785	NUMBER OF HOME ARTICLES : 4
ARTICL=5	0.084178	NUMBER OF HOME ARTICLES : 5
<MODAGE	0.340677	< MODAL AGE
>MODAGE	-0.179236	> MODAL AGE
HWK:NONE	0.127687	TIME ON HOMEWORK: NONE DONE
HWK:<1	0.071980	TIME ON HOMEWORK: < 1 HOUR
HWK:1-2	0.171361	TIME ON HOMEWORK: 1-2 HOURS
HWK:>2	0.106884	TIME ON HOMEWORK: > 2 HOURS
HWK:MISS	-0.207474	TIME ON HOMEWORK: MISSING
NON_PUBL	-0.050240	PRIVATE, CATHOLIC, BIA, DOD MISSING
TV-LINER	-0.039059	HOURS OF TV WATCH (LINEAR) 0-6
TV-QUAD	-0.001544	HOURS OF IV WATCH (QUAD) 0-36
MOTH_WKS	-0.059821	MOTHER WORKS OUTSIDE HOME
LANG_MIN	-0.258029	LANGUAGE MINORITY
LANG_MIS	-0.444430	LANGUAGE MINORITY: MISSING
%LN1-5	-0.018904	PERCENT IN LUNCH PROGRAM: 1-5%
%LN6-10	0.095479	PERCENT IN LUNCH PROGRAM: 6-10%
%LN11-15	-0.167179	PERCENT IN LUNCH PROGRAM: 11-15%
%LN16-50	-0.145792	PERCENT IN LUNCH PROGRAM: 16-50%
%LN51-75	-0.212956	PERCENT IN LUNCH PROGRAM: 51-75%
%LN76-90	-0.812692	PERCENT IN LUNCH PROGRAM: 76-90%
%LN>90	-0.321054	PERCENT IN LUNCH PROGRAM: OVER 90%
INTEGSCH	-0.148655	INTEGRATED SCHOOL (50-79.9% WHITE)
WHIT_MIS	-0.215399	WHITE (80-100%) OR MISSING
GRINSCHL	0.365151	GRADES IN SCHOOL
11+PAGES	0.690277	PAGES READ: 11 + PAGES
6-10PAGE	0.635034	PAGES READ: 6-10 PAGES
1-5PAGES	0.470425	PAGES READ: 1-5 PAGES
#REPORTS	-0.035044	NUMBER OF ESSAYS, REPORTS PRODUCED
<MODGRD	-0.439129	< MODAL GRADE
>MODGRD	-0.226734	> MODAL GRADE

Table F-45
Estimated Effects for Writing Trend Conditioning Variable Contrasts, Age 17

Contrast	Estimated Effect	Description
OVERALL	-1.560249	OVERALL
FEMALE	0.331904	GENDER: FEMALE
WHIT/OTH	0.419154	ETHNICITY: EVERYONE EXCL BLK, HISP
HISPANIC	0.238586	ETHNICITY: HISPANIC
HI_METRO	0.005146	STOC: HIGH METROPOLITAN
STOC-OTH	-0.012870	STOC: OTHER
S_EAST	-0.104561	REGION: SOUTHEAST
CENTRAL	-0.030054	REGION: CENTRAL
WEST	0.080973	REGION: WEST
HS_GRAD	0.010072	PARED: HIGH SCHOOL GRADUATE
POST_HS	0.204556	PARED: POST HIGH SCHOOL
COL_GRAD	0.223519	PARED: COLLEGE GRADUATE
PARED-?	-0.172507	PARED: MISSING, I DON'T KNOW
GR08/A13	0.000000	GRADE 08 / AGE 13
GR11/A17	0.000000	GRADE 11 / AGE 17
ARTICL=4	0.175156	NUMBER OF HOME ARTICLES : 4
ARTICL=5	0.174609	NUMBER OF HOME ARTICLES : 5
<MODAGE	0.213968	< MODAL AGE
>MODAGE	-0.217240	> MODAL AGE
HWK:NONE	0.145861	TIME ON HOMEWORK: NONE DONE
HWK:<1	0.088279	TIME ON HOMEWORK: < 1 HOUR
HWK:1-2	0.252817	TIME ON HOMEWORK: 1-2 HOURS
HWK:>2	0.115699	TIME ON HOMEWORK: > 2 HOURS
HWK:MISS	-1.119650	TIME ON HOMEWORK: MISSING
NON_PUBL	-0.133530	PRIVATE, CATHOLIC, BIA, DOD MISSING
TV-LINER	-0.005998	HOURS OF TV WATCH (LINEAR) 0-6
TV-QUAD	-0.008268	HOURS OF IV WATCH (QUAD) 0-36
MOTH_WKS	-0.031037	MOTHER WORKS OUTSIDE HOME
LANG_MIN	-0.383967	LANGUAGE MINORITY
LANG_MIS	-0.508163	LANGUAGE MINORITY: MISSING
%LN1-5	-0.131312	PERCENT IN LUNCH PROGRAM: 1-5%
%LN6-10	-0.079741	PERCENT IN LUNCH PROGRAM: 6-10%
%LN11-15	-0.108677	PERCENT IN LUNCH PROGRAM: 11-15%
%LN16-50	-0.182744	PERCENT IN LUNCH PROGRAM: 16-50%
%LN51-75	-0.069992	PERCENT IN LUNCH PROGRAM: 51-75%
%LN76-90	0.339175	PERCENT IN LUNCH PROGRAM: 76-90%
%LN>90	-0.051523	PERCENT IN LUNCH PROGRAM: OVER 90%
INTEGSCH	0.018101	INTEGRATED SCHOOL (50-79.9% WHITE)
WHIT_MIS	0.131794	WHITE (80-100%) OR MISSING
GRINSCHL	0.241801	GRADES IN SCHOOL
11+PAGES	0.189602	PAGES READ: 11 + PAGES
6-10PAGE	0.088514	PAGES READ: 6-10 PAGES
1-5PAGES	-0.100068	PAGES READ: 1-5 PAGES
#REPORTS	-0.033460	NUMBER OF ESSAYS, REPORTS PRODUCED
<MODGRD	-0.290916	< MODAL GRADE
>MODGRD	0.275791	> MODAL GRADE

Table F-46
Estimated Effects for Science Trend Conditioning Variable Contrasts, Age 9

Contrast	Estimated Effect	Description
OVERALL	0.237701	OVERALL CONSTANT
GENDER2	-0.256680	SEX (FEMALE)
ETHNIC2	-0.852503	OBSERVED ETHNICITY (BLACK)
ETHNIC3	-0.876387	OBSERVED ETHNICITY (HISPANIC)
ETHNIC4	-0.294888	OBSERVED ETHNICITY (ASIAN)
STOC2	-0.290194	SIZE AND TYPE OF COMMUNITY (LO METRO)
STOC3	0.381629	SIZE AND TYPE OF COMMUNITY (HI METRO)
REGION2	-0.047785	REGION (SOUTHEAST)
REGION3	0.097374	REGION (CENTRAL)
REGION4	-0.100288	REGION (WEST)
PARED2	-0.013854	PARENTS EDUCATION (HIGH SCHOOL GRAD)
PARED3	0.270289	PARENTS EDUCATION (POST HIGH SCHOOL)
PARED4	0.244520	PARENTS EDUCATION (COLLEGE GRAD)
PARED_	-0.011398	PARENTS EDUCATION (MISSING, I DON'T KNOW
MODGRD 1	-0.661367	< MODAL GRADE
MODGRD 2	0.314743	> MODAL GRADE
HOMEITM2	0.227518	3 ITEMS IN THE HOME
HOMEITM3	0.363637	4 ITEMS IN THE HOME
RAC/SEX1	-0.068697	BLACK, FEMALE
RAC/SEX2	0.298483	HISPANIC, FEMALE
RAC/SEX3	0.158868	ASIAN AMERICAN, FEMALE
RAC/PED1	0.037640	BLACK, HIGH SCHOOL GRAD
RAC/PED2	-0.312680	BLACK, POST HIGH SCHOOL
RAC/PED3	-0.022464	BLACK, COLLEGE GRAD
RAC/PED4	0.149869	BLACK, MISSING
RAC/PED5	0.308909	HISPANIC, HIGH SCHOOL GRAD
RAC/PED6	0.023040	HISPANIC, POST HIGH SCHOOL
RAC/PED7	0.066675	HISPANIC, COLLEGE GRAD
RAC/PED8	0.103125	HISPANIC, MISSING
RAC/PED9	-1.520355	ASIAN AMERICAN, HIGH SCHOOL GRAD
RAC/PE10	-0.557077	ASIAN AMERICAN, POST HIGH SCHOOL
RAC/PE11	-0.126000	ASIAN AMERICAN, COLLEGE GRAD
RAC/PE12	-0.540003	ASIAN AMERICAN, MISSING
SCHTYPE	-0.085645	SCHOOL (NONPUBLIC)
LANGHM1	-0.008109	OTHER THAN ENGLISH AT HOME (SOMETIMES)
LANGHM2	-0.574342	OTHER THAN ENGLISH AT HOME (ALWAYS)
RAC/LNG1	-0.037137	BLACK, ALWAYS OTHER THAN ENGLISH
RAC/LNG2	0.017083	BLACK, SOMETIMES OTHER THAN ENGLISH
RAC/LNG3	0.659214	HISPANIC, ALWAYS OTHER THAN ENGLISH
RAC/LNG4	0.141370	HISPANIC, SOMETIMES OTHER THAN ENGLISH
RAC/LNG5	0.717662	ASIAN AM, ALWAYS OTHER THAN ENGLISH
RAC/LNG6	0.111598	ASIAN AM, SOMETIMES OTHER THAN ENGLISH
DRACE2	-0.029279	DERIVED RACE (BLACK)
DRACE3	-0.346203	DERIVED RACE (HISPANIC)
DRACE4	0.222890	DERIVED RACE (ASIAN AMERICAN)

Table F-47
Estimated Effects for Science Trend Conditioning Variable Contrasts, Age 13

Contrast	Estimated Effect	Description
OVERALL	-0.155237	OVERALL CONSTANT
GENDER2	-0.214546	SEX (FEMALE)
ETHNIC2	-0.932830	OBSERVED ETHNICITY (BLACK)
ETHNIC3	0.175256	OBSERVED ETHNICITY (HISPANIC)
ETHNIC4	-0.952854	OBSERVED ETHNICITY (ASIAN)
STOC2	-0.366036	SIZE AND TYPE OF COMMUNITY (LO METRO)
STOC3	0.165310	SIZE AND TYPE OF COMMUNITY (HI METRO)
REGION2	0.128002	REGION (SOUTHEAST)
REGION3	0.137491	REGION (CENTRAL)
REGION4	0.058191	REGION (WEST)
PARED2	0.179317	PARENTS EDUCATION (HIGH SCHOOL GRAD)
PARED3	0.433114	PARENTS EDUCATION (POST HIGH SCHOOL)
PARED4	0.510220	PARENTS EDUCATION (COLLEGE GRAD)
PARED_	-0.119328	PARENTS EDUCATION (MISSING, I DON'T KNOW
MODGRD 1	-0.393358	< MODAL GRADE
MODGRD 2	0.830102	> MODAL GRADE
HOMEITM2	0.241793	3 ITEMS IN THE HOME
HOMEITM3	0.313230	4 ITEMS IN THE HOME
RAC/SEX1	0.230202	BLACK, FEMALE
RAC/SEX2	0.145062	HISPANIC, FEMALE
RAC/SEX3	0.012632	ASIAN AMERICAN, FEMALE
RAC/PED1	0.087895	BLACK, HIGH SCHOOL GRAD
RAC/PED2	0.528668	BLACK, POST HIGH SCHOOL
RAC/PED3	0.240478	BLACK, COLLEGE GRAD
RAC/PED4	0.422786	BLACK, MISSING
RAC/PED5	-0.050704	HISPANIC, HIGH SCHOOL GRAD
RAC/PED6	0.088612	HISPANIC, POST HIGH SCHOOL
RAC/PED7	-0.249976	HISPANIC, COLLEGE GRAD
RAC/PED8	-0.041886	HISPANIC, MISSING
RAC/PED9	0.798376	ASIAN AMERICAN, HIGH SCHOOL GRAD
RAC/PE10	0.581394	ASIAN AMERICAN, POST HIGH SCHOOL
RAC/PE11	0.812608	ASIAN AMERICAN, COLLEGE GRAD
RAC/PE12	0.746130	ASIAN AMERICAN, MISSING
SCHTYPE	-0.122572	SCHOOL (NONPUBLIC)
LANGHM1	0.122135	OTHER THAN ENGLISH AT HOME (SOMETIMES)
LANGHM2	-0.155915	OTHER THAN ENGLISH AT HOME (ALWAYS)
RAC/LNG1	-0.306472	BLACK, ALWAYS OTHER THAN ENGLISH
RAC/LNG2	0.024077	BLACK, SOMETIMES OTHER THAN ENGLISH
RAC/LNG3	-0.158028	HISPANIC, ALWAYS OTHER THAN ENGLISH
RAC/LNG4	-0.213561	HISPANIC, SOMETIMES OTHER THAN ENGLISH
RAC/LNG5	0.206004	ASIAN AM, ALWAYS OTHER THAN ENGLISH
RAC/LNG6	0.167182	ASIAN AM, SOMETIMES OTHER THAN ENGLISH
DRACE2	-0.215731	DERIVED RACE (BLACK)
DRACE3	-0.551525	DERIVED RACE (HISPANIC)

Table F-47 (continued)
Estimated Effects for Science Trend Conditioning Variable Contrasts, Age 13

Contrast	Estimated Effect	Description
DRACE4	0.047742	DERIVED RACE (ASIAN AMERICAN)
HW-NO	-0.036662	HOMEWORK (NONE ASSIGNED)
HW-YES	-0.164016	HOMEWORK (YES - DIDN'T DO)
HW-234	0.063323	HOMEWORK (1\2 HR TO 2 HOURS)

Table F-48
Estimated Effects for Science Trend Conditioning Variable Contrasts, Age 17

Contrast	Estimated Effect	Description
OVERALL	-0.413756	OVERALL CONSTANT
GENDER2	-0.335942	SEX (FEMALE)
ETHNIC2	-0.585301	OBSERVED ETHNICITY (BLACK)
ETHNIC3	-0.150576	OBSERVED ETHNICITY (HISPANIC)
ETHNIC4	-0.073656	OBSERVED ETHNICITY (ASIAN)
STOC2	-0.167269	SIZE AND TYPE OF COMMUNITY (LO METRO)
STOC3	-0.096432	SIZE AND TYPE OF COMMUNITY (HI METRO)
REGION2	0.023967	REGION (SOUTHEAST)
REGION3	0.136414	REGION (CENTRAL)
REGION4	0.066085	REGION (WEST)
PARED2	0.050717	PARENTS EDUCATION (HIGH SCHOOL GRAD)
PARED3	0.204889	PARENTS EDUCATION (POST HIGH SCHOOL)
PARED4	0.275781	PARENTS EDUCATION (COLLEGE GRAD)
PARED_	-0.169730	PARENTS EDUCATION (MISSING, I DON'T KNOW
MODGRD 1	-0.311114	< MODAL GRADE
MODGRD 2	-0.005602	> MODAL GRADE
HOMEITM2	0.224909	3 ITEMS IN THE HOME
HOMEITM3	0.215292	4 ITEMS IN THE HOME
RAC/SEX1	0.137868	BLACK, FEMALE
RAC/SEX2	-0.025479	HISPANIC, FEMALE
RAC/SEX3	0.035208	ASIAN AMERICAN, FEMALE
RAC/PED1	0.018093	BLACK, HIGH SCHOOL GRAD
RAC/PED2	0.014291	BLACK, POST HIGH SCHOOL
RAC/PED3	-0.108222	BLACK, COLLEGE GRAD
RAC/PED4	0.154685	BLACK, MISSING
RAC/PED5	-0.016733	HISPANIC, HIGH SCHOOL GRAD
RAC/PED6	0.238679	HISPANIC, POST HIGH SCHOOL
RAC/PED7	-0.159268	HISPANIC, COLLEGE GRAD
RAC/PED8	-0.407308	HISPANIC, MISSING
RAC/PED9	0.331197	ASIAN AMERICAN, HIGH SCHOOL GRAD
RAC/PE10	0.037220	ASIAN AMERICAN, POST HIGH SCHOOL
RAC/PE11	0.333585	ASIAN AMERICAN, COLLEGE GRAD
RAC/PE12	0.191232	ASIAN AMERICAN, MISSING
SCHTYPE	-0.002745	SCHOOL (NONPUBLIC)
LANGHM1	0.006612	OTHER THAN ENGLISH AT HOME (SOMETIMES)
LANGHM2	-0.287407	OTHER THAN ENGLISH AT HOME (ALWAYS)
RAC/LNG1	0.438359	BLACK, ALWAYS OTHER THAN ENGLISH
RAC/LNG2	-0.113115	BLACK, SOMETIMES OTHER THAN ENGLISH
RAC/LNG3	0.420432	HISPANIC, ALWAYS OTHER THAN ENGLISH
RAC/LNG4	0.139094	HISPANIC, SOMETIMES OTHER THAN ENGLISH
RAC/LNG5	-0.239001	ASIAN AM, ALWAYS OTHER THAN ENGLISH
RAC/LNG6	-0.407667	ASIAN AM, SOMETIMES OTHER THAN ENGLISH
DRACE2	-0.206510	DERIVED RACE (BLACK)
DRACE3	-0.345286	DERIVED RACE (HISPANIC)

Table F-48 (continued)
Estimated Effects for Science Trend Conditioning Variable Contrasts, Age 17

Contrast	Estimated Effect	Description
DRACE4	-0.012900	DERIVED RACE (ASIAN AMERICAN)
HW-NO	-0.250457	HOMEWORK (NONE ASSIGNED)
HW-YES	-0.097001	HOMEWORK (YES - DIDN'T DO)
HW-234	0.006259	HOMEWORK (1\2 HR TO 2 HOURS)
NSCI-GS	-0.112945	NSCI (GENERAL SCIENCE)
NSCI-BI	0.139804	NSCI (BIOLOGY)
NSCI-CH	0.649440	NSCI (CHEMISTRY)
NSCI-PH	0.834929	NSCI (PHYSICS)
HSPROG1	0.358359	HS PROGRAM (COLLEGE PREP)
HSPROG2	-0.036378	HS PROGRAM (VOCAT/TECHN)

Table F-49
Proportion of Variance of the Conditioning Variable Contrasts Accounted for by the Principal Components Used in the Conditioning Model for Reading Main Conditioning Variables, Age 9/Grade 4

Contrast	Proportion of Variance	Contrast	Proportion of Variance	Contrast	Proportion of Variance
FEMALE	0.968411	CMP4SCH3	0.968220	ALOUD-S2	0.896899
BLACK	0.805821	CMP4SCH4	0.939574	ALOUD-S3	0.913193
HISPANIC	0.862039	CMP4SCH5	0.989909	ALOUD-S4	0.895456
ASIAN	0.952039	SLP_RD-L	0.833203	ALOUD-S?	0.708818
HI_METRO	0.891152	NBOOKS-2	0.958739	SILNT-S2	0.935492
STOC-OTH	0.904799	NBOOKS-3	0.965212	SILNT-S3	0.951808
S_EAST	0.913381	NBOOKS-4	0.950056	SILNT-S4	0.952124
CENTRAL	0.904380	NBOOKS-?	0.841459	SILNT-S?	0.728344
WEST	0.902477	READ-GD	0.858655	RDLOG-S2	0.925260
HS_GRAD	0.986847	READ-AVG	0.853286	RDLOG-S3	0.924894
POST_HS	0.985704	READ-PR	0.938094	RDLOG-S4	0.893674
COL_GRAD	0.974757	READ-?	0.889602	RDLOG-S?	0.877091
PARED-?	0.974624	RD4FUN-2	0.886767	OWNBK-S2	0.868370
HITEM=3	0.850194	RD4FUN-3	0.894624	OWNBK-S3	0.887782
HITEM=4	0.798097	RD4FUN-4	0.873605	OWNBK-S4	0.897453
TV-LIN1	0.986444	RD4FUN-?	0.774821	OWNBK-S?	0.866962
TV-QUAD1	0.986551	TALKRD-2	0.901400	#QUEST+2	0.873190
HL-SM/AL	0.934608	TALKRD-3	0.919223	#QUEST+3	0.868702
HW4-NONE	0.993354	TALKRD-4	0.886375	#QUEST+4	0.906437
HW4-YES	0.993735	TALKRD-?	0.784530	TESTDIF2	0.915574
HW4-LIN1	0.986902	USELIB-2	0.915972	TESTDIF3	0.899711
HW4QUAD1	0.986414	USELIB-3	0.932383	TESTDIF4	0.917066
MINORITY	0.836124	USELIB-4	0.941362	TESTDIF?	0.904746
INTEGRAT	0.856801	USELIB-?	0.870461	TESTEFF2	0.882760
=MA/<MG	0.821266	VOCAB-S2	0.874123	TESTEFF3	0.850874
=MA/=MG	0.951194	VOCAB-S3	0.915463	TESTEFF4	0.913553
=MA/>MG	0.995417	VOCAB-S4	0.884890	TESTEFF?	0.909955
>MA/=MG	0.953408	VOCAB-S?	0.836347	TESTIMP2	0.941815
NON_PUBL	0.865265	TLKRD-S2	0.921008	TESTIMP3	0.955184
2PARENTS	0.793270	TLKRD-S3	0.935549	TESTIMP4	0.966820
MOM@HM-Y	0.803368	TLKRD-S4	0.910049	TESTIMP?	0.937283
>=6_PGS	0.814136	TLKRD-S?	0.769075	DSOLUTN2	0.868254
>=11_PGS	0.818498	WB/WS-S2	0.891556	DSOLUTN3	0.865638
PRESCH-Y	0.980443	WB/WS-S3	0.909801	DSOLUTN4	0.877534
USA-NO/?	0.974807	WB/WS-S4	0.907561	DSOLUTN5	0.941509
CHGSCH=1	0.889597	WB/WS-S?	0.720666	RPRIOR-N	0.888937
CHGSCH=2	0.872057	WRTRD-S2	0.911198	RPRIOR-?	0.824506
CHGSCH3+	0.825063	WRTRD-S3	0.908225	WPRIOR-N	0.857936
ST4GRD12	0.868206	WRTRD-S4	0.917551	WPRIOR-?	0.841475
ST4GRD3+	0.873276	WRTRD-S?	0.732837	%SUBLUN2	0.877609
DIS@HOM2	0.897096	RDPRJ-S2	0.941327	%SUBLUN3	0.908924
DIS@HOM3	0.928589	RDPRJ-S3	0.942607	%SUBLUN4	0.943898
DIS@HOM4	0.831864	RDPRJ-S4	0.934939	%SUBLUN5	0.872838
CMP4SCH2	0.953347	RDPRJ-S?	0.744449	%SUBLUN?	0.835202

Table F-49 (continued)
Proportion of Variance of the Conditioning Variable Contrasts Accounted for by the Principal Components Used in the Conditioning Model for Reading Main Conditioning Variables, Age 9/Grade 4

Contrast	Proportion of Variance	Contrast	Proportion of Variance	Contrast	Proportion of Variance
%REMRED2	0.875647	TCERT-HI	0.916705	INSTIME3	0.884433
%REMRED3	0.901036	TCERT-?	0.911981	INSTIME?	0.995125
%REMRED4	0.887713	CERTG-N	0.919533	INSGRP-2	0.827286
%REMRED?	0.849516	CERTG-?	0.962512	INSGRP-3	0.847677
%ENR/YR2	0.905148	BACHELRS	0.948985	INSGRP-4	0.925345
%ENR/YR3	0.924587	MASTERS	0.952137	TRADE_RM	0.869615
%ENR/YR4	0.911493	SPECLIST	0.925196	BAS&TRA	0.852453
%ENR/YR?	0.872284	DOCTORAT	0.942788	OTHER_RM	0.885565
DEPTLIZD	0.866424	DEGREE-?	0.991094	RMATS-?	0.944792
REGRPED	0.844520	UGR_ED-Y	0.827262	NEWMAG-2	0.856849
ORGANIZ?	0.811119	UGR_RD-Y	0.843546	NEWMAG-3	0.844627
ABILTY-N	0.859951	GR_ED-Y	0.849432	NEWMAG-?	0.975550
ABILTY-?	0.841999	GR_RD-Y	0.885562	RDKITS-2	0.884184
RD_POL-N	0.870119	GR_NO-Y	0.884929	RDKITS-3	0.835030
RD_POL-?	0.855761	READCS-N	0.816619	RDKITS-?	0.952294
WR_POL-N	0.874921	READCS-?	0.977015	SOFTWA-2	0.876926
WR_POL-?	0.843663	TCHAIDE2	0.899179	SOFTWA-3	0.857186
PARAID-O	0.865760	TCHAIDE3	0.910426	SOFTWA-?	0.962437
PARAID-N	0.826476	TCHAIDE4	0.909115	VARBKS-2	0.860561
PARAID-?	0.845769	TCHAIDE5	0.912936	VARBKS-3	0.848940
RESOURC2	0.913534	TCHAIDE?	0.987454	VARBKS-?	0.976944
RESOURC3	0.928506	T_NCLAS2	0.843182	OTHMAT-2	0.885636
RESOURC4	0.955071	T_NCLAS3	0.873845	OTHMAT-3	0.869329
RESOURC?	0.985601	T_NCLAS4	0.865225	OTHMAT-?	0.975225
TMCH-PAR	0.934690	T_NCLAS5	0.878936	VOCAB-T2	0.774839
TMCH-COM	0.991769	T_NCLAS?	0.892907	VOCAB-T3	0.922527
T_FEMALE	0.844452	INSERVR2	0.929725	VOCAB-T4	0.868127
T_SEX-?	0.985007	INSERVR3	0.950438	VOCAB-T?	0.992206
T_BLACK	0.887979	INSERVR4	0.926976	ALOUD-T2	0.824185
T_HISP	0.979618	INSERVR5	0.936416	ALOUD-T3	0.869815
T_ASIAN	0.976326	INSERVR?	0.986770	ALOUD-T4	0.831030
T_AM.IND	0.948348	CERTR-N	0.864924	ALOUD-T?	0.996447
T_RACE-?	0.970334	CERTR-NS	0.816650	TLKRD-T2	0.846830
T_MEXICN	0.976691	CERTR-?	0.925135	TLKRD-T3	0.870172
T_PUERTO	0.942130	CERTE-N	0.894310	TLKRD-T4	0.850706
T_CUBAN	0.975530	CERTE-NS	0.877478	TLKRD-T?	0.995903
T_OTHER	0.961370	CERTE-?	0.928634	WRTRD-T2	0.859078
T_HISP-?	0.706660	AB_CLA-N	0.815324	WRTRD-T3	0.897068
T_YREXP2	0.915139	AB_CLA-?	0.994344	WRTRD-T4	0.843006
T_YREXP3	0.950204	AB_READ2	0.927867	WRTRD-T?	0.994804
T_YREXP4	0.952345	AB_READ3	0.903468	WB/WS-T2	0.847460
T_YREXP5	0.949425	AB_READ4	0.910651	WB/WS-T3	0.906428
T_YREXP?	0.988118	AB_READ?	0.987122	WB/WS-T4	0.818457
TCERT-RG	0.923488	INSTIME2	0.842013	SILNT-T2	0.830671

Table F-49 (continued)
Proportion of Variance of the Conditioning Variable Contrasts Accounted for by the Principal Components Used in the Conditioning Model for Reading Main Conditioning Variables, Age 9/Grade 4

Contrast	Proportion of Variance	Contrast	Proportion of Variance	Contrast	Proportion of Variance
SILNT-T3	0.872312	EMPCON-?	0.980577	TREMP1-2	0.931635
SILNT-T4	0.960957	MCTEST-2	0.888298	TREMP2-2	0.863566
SILNT-T?	0.994705	MCTEST-3	0.902293	TREMP3-2	0.885611
OWNBK-T2	0.845726	MCTEST-?	0.992775		
OWNBK-T3	0.897733	SATEST-2	0.846290		
OWNBK-T4	0.833031	SATEST-3	0.836695		
RDLOG-T2	0.867172	SATEST-?	0.994707		
RDLOG-T3	0.905594	WRTPAR-2	0.833673		
RDLOG-T4	0.870854	WRTPAR-3	0.858272		
RDLOG-T?	0.992738	WRTPAR-?	0.990359		
RDPRJ-T2	0.916333	OBSERV-2	0.839912		
RDPRJ-T3	0.932028	OBSERV-3	0.815611		
RDPRJ-T4	0.854117	OBSERV-?	0.981035		
RDPRJ-T?	0.992365	ORLTST-2	0.828951		
DECODE-2	0.884250	ORLTST-3	0.856347		
DECODE-3	0.881524	ORLTST-?	0.994687		
DECODE-?	0.982763	PROJCT-2	0.894335		
ORALRD-2	0.841808	PROJCT-3	0.911378		
ORALRD-3	0.844357	PROJCT-?	0.989410		
ORALRD-?	0.993839	RDPORT-2	0.907085		
VOCABY-2	0.854388	RDPORT-3	0.894751		
VOCABY-3	0.859505	RDPORT-?	0.973701		
VOCABY-?	0.994507	SLFREP-2	0.918756		
CMPREH-2	0.853327	SLFREP-3	0.924216		
STRATG-2	0.822833	SLFREP-?	0.984732		
STRATG-3	0.915115	CLALIB-2	0.930338		
STRATG-?	0.991514	CLALIB-3	0.909787		
EMPPHO-M	0.912156	CLALIB-4	0.865698		
EMPPHO-L	0.917981	CLALIB-5	0.987538		
EMPPHO-?	0.987960	CLALIB-2	0.890954		
EMPINT-M	0.834838	CLALIB-3	0.908391		
EMPINT-L	0.869129	CLALIB-4	0.873484		
EMPINT-?	0.995048	CLALIB-5	0.991090		
EMPLAN-M	0.866137	COMP-DIF	0.876076		
EMPLAN-L	0.879627	COMP-AVL	0.897722		
EMPLAN-?	0.977216	COMP-?	0.986478		
EMPLIT-M	0.854918	MATH-NO	0.820753		
EMPLIT-L	0.879181	MATH-NS	0.873816		
EMPLIT-?	0.990980	M/OTH-?	0.858445		
EMPCON-M	0.857681	UMAJO-N	0.797345		
EMPCON-L	0.830368	GMAJO-N	0.776875		
EMPCON-?	0.991808	TRNS4-6	0.901459		
EMPCON-M	0.923511	TRNS7	0.922068		
EMPCON-L	0.912093	TRNS8	0.923112		

Table F-50
Proportion of Variance of the Conditioning Variable Contrasts Accounted for by the Principal Components Used in the Conditioning Model for Reading Main Conditioning Variables, Age 13/Grade 8

Contrast	Proportion of Variance	Contrast	Proportion of Variance	Contrast	Proportion of Variance
FEMALE	0.966601	CMP4SCH5	0.796760	TLKRD-S?	0.766376
BLACK	0.834509	ST8GRD12	0.961201	WB/WS-S2	0.899301
HISPANIC	0.800865	ST8GRD35	0.947691	WB/WS-S3	0.910270
ASIAN	0.923206	ST8GRD>5	0.883236	WB/WS-S4	0.901284
HI_METRO	0.894540	SLP_RD-L	0.810910	WB/WS-S?	0.701845
STOC-OTH	0.874193	NBOOKS-2	0.918195	WRTRD-S2	0.935773
S_EAST	0.883690	NBOOKS-3	0.933196	WRTRD-S3	0.921875
CENTRAL	0.908854	NBOOKS-4	0.941532	WRTRD-S4	0.944221
WEST	0.895775	NBOOKS-?	0.961257	WRTRD-S?	0.746351
HS_GRAD	0.967441	READ-GD	0.898578	RDPRJ-S2	0.969826
POST_HS	0.974365	READ-AVG	0.883700	RDPRJ-S3	0.961700
COL_GRAD	0.958499	READ-PR	0.956865	RDPRJ-S4	0.962046
PARED-?	0.976735	READ-?	0.962119	RDPRJ-S?	0.727420
HITEM=3	0.877102	RD4FUN-2	0.948752	ALOUD-S2	0.884839
HITEM=4	0.848539	RD4FUN-3	0.930279	ALOUD-S3	0.874852
TV-LIN1	0.985110	RD4FUN-4	0.925858	ALOUD-S4	0.893204
TV-QUAD1	0.984912	RD4FUN-?	0.871491	ALOUD-S?	0.608110
HL-SM/AL	0.874945	RDNOVL-2	0.953471	SILNT-S2	0.879908
HWC-NONE	0.965908	RDNOVL-3	0.949909	SILNT-S3	0.871403
HWC-YES	0.967219	RDNOVL-4	0.948961	SILNT-S4	0.915466
HW-LIN1	0.978331	RDNOVL-?	0.784554	SILNT-S?	0.703296
HW-QUAD1	0.978937	RDNEWP-2	0.894673	RDLOG-S2	0.941702
MINORITY	0.838093	RDNEWP-3	0.911463	RDLOG-S3	0.937496
INTEGRAT	0.863192	RDNEWP-4	0.891702	RDLOG-S4	0.911634
=MA/<MG	0.994162	RDNEWP-?	0.681282	RDLOG-S?	0.802001
=MA/=MG	0.993667	RDMAGZ-2	0.875881	OWNBK-S2	0.935849
=MA/>MG	0.968676	RDMAGZ-3	0.874486	OWNBK-S3	0.932632
>MA/=MG	0.995022	RDMAGZ-4	0.921729	OWNBK-S4	0.919755
NON_PUBL	0.843024	RDMAGZ-?	0.757444	OWNBK-S?	0.835263
2PARENTS	0.690287	TALKRD-2	0.941136	EXPLA-2	0.949950
MOM@HM-Y	0.836928	TALKRD-3	0.948516	EXPLA-3	0.937160
>=6_PGS	0.783927	TALKRD-4	0.941002	EXPLA-4	0.929097
>=11_PGS	0.786241	TALKRD-?	0.772746	EXPLA-?	0.858746
<=2_DAYS	0.981065	USELIB-2	0.976201	DISCU-2	0.942995
USA-NO/?	0.968273	USELIB-3	0.972469	DISCU-3	0.947434
CHGSCH=1	0.955679	USELIB-4	0.970084	DISCU-4	0.942722
CHGSCH=2	0.969962	USELIB-?	0.885650	DISCU-?	0.864839
CHGSCH3+	0.930705	VOCAB-S2	0.918227	PREDI-2	0.948608
DIS@HOM2	0.904820	VOCAB-S3	0.928782	PREDI-3	0.948397
DIS@HOM3	0.925202	VOCAB-S4	0.911679	PREDI-4	0.942281
DIS@HOM4	0.833185	VOCAB-S?	0.770943	PREDI-?	0.849581
CMP4SCH2	0.960084	TLKRD-S2	0.956669	LIBRS-2	0.972618
CMP4SCH3	0.948232	TLKRD-S3	0.959292	LIBRS-3	0.965283
CMP4SCH4	0.944209	TLKRD-S4	0.944866	LIBRS-4	0.964457

Table F-50 (continued)
Proportion of Variance of the Conditioning Variable Contrasts Accounted for by the Principal Components Used in the Conditioning Model for Reading Main Conditioning Variables, Age 13/Grade 8

Contrast	Proportion of Variance	Contrast	Proportion of Variance	Contrast	Proportion of Variance
LIBRS-?	0.892093	%REMRED4	0.844669		
LIBBR-2	0.972058	%REMRED?	0.841602		
LIBBR-3	0.966857	%ENR/YR2	0.906562		
LIBBR-4	0.962444	%ENR/YR3	0.886347		
LIBBR-?	0.929267	%ENR/YR4	0.872209		
LIBIN-2	0.959817	%ENR/YR?	0.853853		
LIBIN-3	0.959832	SEMIDEPT	0.921483		
LIBIN-4	0.947906	DEPTLIZD	0.936784		
LIBIN-?	0.933886	ORGANIZ?	0.921887		
LIBQU-2	0.922928	ENG/AB-N	0.958314		
LIBQU-3	0.920979	ENG/AB-?	0.892761		
LIBQU-4	0.856135	CMPECL-N	0.805489		
LIBQU-?	0.920014	CMPECL-?	0.865434		
#QUEST+2	0.900729	CMPELB-N	0.944069		
#QUEST+3	0.894511	CMPELB-?	0.913036		
#QUEST+4	0.924223	CMPEBR-N	0.891835		
TESTDIF2	0.968351	CMPEBR-?	0.866182		
TESTDIF3	0.956964	PARAID-O	0.899026		
TESTDIF4	0.960618	PARAID-N	0.892585		
TESTDIF?	0.799130	PARAID-?	0.941902		
TESTEFF2	0.928475				
TESTEFF3	0.918653				
TESTEFF4	0.948340				
TESTEFF?	0.817838				
TESTIMP2	0.920421				
TESTIMP3	0.899537				
TESTIMP4	0.907401				
TESTIMP?	0.873831				
DSOLUTN2	0.840393				
DSOLUTN3	0.851037				
DSOLUTN4	0.906344				
DSOLUTN5	0.838080				
ENG-T=1S	0.723095				
RPRIOR-N	0.801014				
RPRIOR-?	0.897268				
WPRIOR-N	0.818808				
WPRIOR-?	0.830089				
%SUBLUN2	0.815849				
%SUBLUN3	0.942951				
%SUBLUN4	0.910907				
%SUBLUN5	0.918413				
%SUBLUN?	0.884138				
%REMRED2	0.928088				
%REMRED3	0.870843				

Table F-51

Proportion of Variance of the Conditioning Variable Contrasts Accounted for by the Principal Components Used in the Conditioning Model for Reading Main Conditioning Variables, Age 17/Grade 12

Contrast	Proportion of Variance	Contrast	Proportion of Variance	Contrast	Proportion of Variance
FEMALE	0.909746	#LAN-L01	0.754570	TALKRD-3	0.932781
BLACK	0.852308	#VOC-Y	0.902413	TALKRD-4	0.942447
HISPANIC	0.827930	#VOC-L01	0.882447	TALKRD-?	0.869551
ASIAN	0.822495	#ART-Y	0.896538	USELIB-2	0.980400
HI_METRO	0.895324	#ART-L01	0.933482	USELIB-3	0.979788
STOC-OTH	0.897173	USA-NO/?	0.755078	USELIB-4	0.973403
S_EAST	0.878462	DIS@HOM2	0.911006	USELIB-?	0.912740
CENTRAL	0.904903	DIS@HOM3	0.920741	VOCAB-S2	0.915925
WEST	0.911951	DIS@HOM4	0.876286	VOCAB-S3	0.928628
HS_GRAD	0.966956	CMP4SCH2	0.937624	VOCAB-S4	0.927020
POST_HS	0.967719	CMP4SCH3	0.913228	VOCAB-S?	0.838174
COL_GRAD	0.957967	CMP4SCH4	0.894121	TLKRD-S2	0.946519
PARED-?	0.986593	CMP4SCH5	0.733151	TLKRD-S3	0.946328
HITEM=3	0.885724	VOC/BUSI	0.977776	TLKRD-S4	0.929534
HITEM=4	0.867063	2-YR_COL	0.966696	TLKRD-S?	0.830468
TV-LIN1	0.981368	4-YR_COL	0.931544	WB/WS-S2	0.922680
TV-QUAD1	0.980458	MILITARY	0.983440	WB/WS-S3	0.931427
HL-SM/AL	0.813326	OTHERACT	0.899177	WB/WS-S4	0.922030
HWC-NONE	0.981347	SLP_RD-L	0.775304	WB/WS-S?	0.664486
HWC-YES	0.982144	NBOOKS-2	0.882469	WRTRD-S2	0.928660
HW-LIN1	0.977781	NBOOKS-3	0.913234	WRTRD-S3	0.910857
HW-QUAD1	0.977019	NBOOKS-4	0.940458	WRTRD-S4	0.950665
MINORITY	0.834252	NBOOKS-?	0.970200	WRTRD-S?	0.764350
INTEGRAT	0.924272	READ-GD	0.912029	RDPRJ-S2	0.970707
=MA/<MG	0.987859	READ-AVG	0.894508	RDPRJ-S3	0.966160
=MA/=MG	0.990940	READ-PR	0.961864	RDPRJ-S4	0.967820
>MA/=MG	0.991531	READ-?	0.971432	RDPRJ-S?	0.805583
NON_PUBL	0.893317	RD4FUN-2	0.937160	ALOUD-S2	0.915317
2PARENTS	0.804110	RD4FUN-3	0.925677	ALOUD-S3	0.913179
MOM@HM-Y	0.807619	RD4FUN-4	0.924630	ALOUD-S4	0.919696
> =6_PGS	0.806850	RD4FUN-?	0.893085	ALOUD-S?	0.746613
> =11_PGS	0.803952	RDNOVL-2	0.960026	SILNT-S2	0.858312
< =2_DAYS	0.989205	RDNOVL-3	0.948138	SILNT-S3	0.861098
COL_PREP	0.705776	RDNOVL-4	0.954401	SILNT-S4	0.924638
VOC/TECH	0.948792	RDNOVL-?	0.807871	SILNT-S?	0.705951
#ENG-Y	0.875087	RDNEWP-2	0.891668	RDLOG-S2	0.947659
#ENG-L01	0.895325	RDNEWP-3	0.891897	RDLOG-S3	0.941227
#MATH-Y	0.915875	RDNEWP-4	0.884677	RDLOG-S4	0.919396
#MAT-L01	0.902152	RDNEWP-?	0.697740	RDLOG-S?	0.855187
#SCI-Y	0.902027	RDMAGZ-2	0.873669	OWNBK-S2	0.979115
#SCI-L01	0.864876	RDMAGZ-3	0.869113	OWNBK-S3	0.968600
#HIS-Y	0.900333	RDMAGZ-4	0.951450	OWNBK-S4	0.960821
#HIS-L01	0.813890	RDMAGZ-?	0.698435	OWNBK-S?	0.888535
#FLANG-Y	0.903780	TALKRD-2	0.927505	EXPLA-2	0.860405

Table F-51 (continued)
Proportion of Variance of the Conditioning Variable Contrasts Accounted for by the Principal Components Used in the Conditioning Model for Reading Main Conditioning Variables, Age 17/Grade 12

Contrast	Proportion of Variance	Contrast	Proportion of Variance	Contrast	Proportion of Variance
EXPLA-3	0.932526	DSOLUTN4	0.913787		
EXPLA-4	0.807665	DSOLUTN5	0.887613		
EXPLA-?	0.846683	%SUBLUN2	0.894361		
DISCU-2	0.857964	%SUBLUN3	0.861853		
DISCU-3	0.936312	%SUBLUN4	0.909295		
DISCU-4	0.798818	%SUBLUN5	0.919914		
DISCU-?	0.859133	%SUBLUN?	0.784679		
PREDI-2	0.949353	%REMRED2	0.943571		
PREDI-3	0.953967	%REMRED3	0.904774		
PREDI-4	0.942777	%REMRED4	0.807176		
PREDI-?	0.844880	%REMRED?	0.918769		
LIBRS-2	0.845235	%ENR/YR2	0.878837		
LIBRS-3	0.967087	%ENR/YR3	0.911722		
LIBRS-4	0.881873	%ENR/YR4	0.879769		
LIBRS-?	0.887103	%ENR/YR?	0.889740		
LIBBR-2	0.830922	ENG/AB-N	0.863987		
LIBBR-3	0.973301	ENG/AB-?	0.970941		
LIBBR-4	0.876575	#SEMENG8	0.749141		
LIBBR-?	0.922191	MAT/AB-N	0.859112		
LIBIN-2	0.971869	MAT/AB-?	0.906581		
LIBIN-3	0.968240	CMPECL-N	0.847581		
LIBIN-4	0.959402	CMPECL-?	0.886883		
LIBIN-?	0.916890	CMPELB-N	0.941905		
LIBQU-2	0.930854	CMPELB-?	0.927489		
LIBQU-3	0.925848	CMPEBR-N	0.837176		
LIBQU-4	0.871059	CMPEBR-?	0.826704		
LIBQU-?	0.897899	AP_ENGL2	0.966796		
#QUEST+2	0.929109	AP_ENGL3	0.965534		
#QUEST+3	0.909477	AP_ENGL4	0.902504		
#QUEST+4	0.930406	AP_ENGL5	0.908311		
TESTDIF2	0.980869	AP_ENGL6	0.905339		
TESTDIF3	0.972302	AP_ENGL?	0.904688		
TESTDIF4	0.970899	PARAID-O	0.945879		
TESTDIF?	0.721811	PARAID-N	0.953531		
TESTEFF2	0.983167	ENGL=45M	0.980758		
TESTEFF3	0.968850	ENGL=60M	0.983802		
TESTEFF4	0.972822	ENGL>90M	0.938034		
TESTEFF?	0.755939				
TESTIMP2	0.961388				
TESTIMP3	0.950042				
TESTIMP4	0.947557				
TESTIMP?	0.903776				
DSOLUTN2	0.861023				
DSOLUTN3	0.864487				

Table F-52
Proportion of Variance of the Conditioning Variable Contrasts Accounted for by the Principal Components Used in the Conditioning Model for Mathematics Main Conditioning Variables, Age 9/Grade 4

Contrast	Proportion of Variance	Contrast	Proportion of Variance	Contrast	Proportion of Variance
FEMALE	0.962948	CMP4SCH3	0.966236	GDMATH-D	0.827275
BLACK	0.781001	CMP4SCH4	0.942436	USMATH-U	0.915480
HISPANIC	0.843402	CMP4SCH5	0.983817	USMATH-D	0.722637
ASIAN	0.954269	SLP_MA-L	0.720731	MA4BOY-U	0.869694
HI_METRO	0.906518	S_TXTBK2	0.879380	MA4BOY-D	0.840823
STOC-OTH	0.904534	S_TXTBK3	0.936089	MATMEM-U	0.849050
S_EAST	0.897840	S_TXTBK4	0.856076	MATMEM-D	0.737294
CENTRAL	0.883641	S_TXTBK?	0.759186	MATJOB-U	0.937065
WEST	0.895983	S_WRKSH2	0.836669	MATJOB-D	0.774341
HS_GRAD	0.987364	S_WRKSH3	0.900682	MA4PRB-U	0.931373
POST_HS	0.988628	S_WRKSH4	0.860623	MA4PRB-D	0.775689
COL_GRAD	0.973978	S_WRKSH?	0.745361	NSTDMA-U	0.875924
PARED-?	0.973037	S_SMGRP2	0.930733	NSTDMA-D	0.867272
HITEM=3	0.841160	S_SMGRP3	0.938351	#QUEST+2	0.852403
HITEM=4	0.795068	S_SMGRP4	0.915111	#QUEST+3	0.836400
TV-LIN	0.986431	S_SMGRP?	0.850420	#QUEST+4	0.901901
TV-QUAD	0.986490	S_OBJCT2	0.941473	TESTDIF2	0.902845
HL-SM/AL	0.940011	S_OBJCT3	0.941200	TESTDIF3	0.895404
HW4-NONE	0.983865	S_OBJCT4	0.925650	TESTDIF4	0.912637
HW4-YES	0.984457	S_OBJCT?	0.796707	TESTEFF2	0.877924
HW4-LIN	0.982259	S_CALCR2	0.967704	TESTEFF3	0.849613
HW4QUAD	0.973637	S_CALCR3	0.959411	TESTEFF4	0.900593
MINORITY	0.847604	S_CALCR4	0.946117	TESTIMP2	0.944324
INTEGRAT	0.869403	S_CALCR?	0.772789	TESTIMP3	0.958997
=MA/<MG	0.844080	S_CMPTR2	0.944952	TESTIMP4	0.968682
=MA/=MG	0.956125	S_CMPTR3	0.969898	TESTIMP?	0.937108
=MA/>MG	0.989203	S_CMPTR4	0.934083	DSOLUTN2	0.884373
>MA/=MG	0.961307	S_CMPTR?	0.929142	DSOLUTN3	0.905048
NON_PUBL	0.891585	S_MATST2	0.929042	DSOLUTN4	0.898896
2PARENTS	0.801622	S_MATST3	0.929408	RPRIOR-N	0.776561
MOM@HM-Y	0.811892	S_MATST4	0.965213	RPRIOR-?	0.825652
>=6_PGS	0.819036	S_MATST?	0.667871	WPRIOR-N	0.798571
>=11_PGS	0.827721	HVCALC-N	0.964646	WPRIOR-?	0.822744
PRESCH-Y	0.970409	S_MATHW1	0.987622	MPRIOR-N	0.798878
USA-NO/?	0.973352	S_MATHW2	0.987611	MPRIOR-?	0.794330
CHGSCH=1	0.868837	S_MATHW3	0.991817	%SUBLUN2	0.902003
CHGSCH=2	0.883868	S_MATHW4	0.992590	%SUBLUN3	0.924239
CHGSCH3+	0.817835	S_MATHW5	0.991414	%SUBLUN4	0.931849
ST4GRD12	0.876854	S_MATHW6	0.983885	%SUBLUN5	0.943228
ST4GRD3+	0.879410	S_MATHW?	0.964506	%SUBLUN6	0.876403
DIS@HOM2	0.863373	MATHLP-N	0.949314	%SUBLUN7	0.906546
DIS@HOM3	0.911585	LKMATH-U	0.917833	%SUBLUN8	0.847837
DIS@HOM4	0.791895	LKMATH-D	0.831919	%SUBLUN?	0.765154
CMP4SCH2	0.952706	GDMATH-U	0.921443	%REMRED2	0.861224

Table F-52 (continued)
Proportion of Variance of the Conditioning Variable Contrasts Accounted for by the Principal Components Used in the Conditioning Model for Mathematics Main Conditioning Variables, Age 9/Grade 4

Contrast	Proportion of Variance	Contrast	Proportion of Variance	Contrast	Proportion of Variance
%REMRED3	0.891417	T_MATHW4	0.923763	T_MUPRB4	0.895408
%REMRED4	0.874119	T_MATHW5	0.895639	T_MUPRB?	0.996196
%REMRED5	0.844562	T_MATHW6	0.911151	EMP_N&O2	0.840922
%REMRED6	0.915576	T_MATHW?	0.993318	EMP_N&O3	0.917794
%REMRED7	0.865716	T_TXTBK2	0.776724	EMP_N&O?	0.988340
%REMRED8	0.933181	T_TXTBK3	0.868002	EMP_MEA2	0.858251
%REMRED?	0.851298	T_TXTBK4	0.951621	EMP_MEA3	0.891629
%REMMAT2	0.819738	T_TXTBK?	0.993571	EMP_MEA?	0.995467
%REMMAT3	0.825347	T_WRKSH2	0.794780	EMP_GEO2	0.937344
%REMMAT4	0.826217	T_WRKSH3	0.866538	EMP_GEO3	0.955648
%REMMAT5	0.795124	T_WRKSH4	0.928398	EMP_GEO?	0.988210
%REMMAT6	0.897771	T_WRKSH?	0.995117	EMP_DSP2	0.905051
%REMMAT7	0.841666	T_SMGRP2	0.849698	EMP_DSP3	0.921559
%REMMAT?	0.802404	T_SMGRP3	0.910453	EMP_DSP?	0.991298
%ENR/YR2	0.875232	T_SMGRP4	0.868452	EMP_ALG2	0.930673
%ENR/YR3	0.904298	T_SMGRP?	0.994211	EMP_ALG3	0.944473
%ENR/YR4	0.886174	T_OBJCT2	0.909953	EMP_ALG?	0.992637
%ENR/YR?	0.859679	T_OBJCT3	0.951951	EMP_F/C2	0.848285
%RETAIN2	0.855615	T_OBJCT4	0.935170	EMP_F/C?	0.996386
%RETAIN3	0.909737	T_OBJCT?	0.992218	EMP_S/P2	0.850502
%RETAIN4	0.846843	T_CALCR2	0.919141	EMP_R/A2	0.801448
%RETAIN5	0.868475	T_CALCR3	0.955639	EMP_R/A3	0.840939
%RETAIN?	0.803786	T_CALCR4	0.951444	EMP_R/A?	0.989137
%TLEAVE2	0.857771	T_CALCR?	0.992460	EMP_CMI2	0.843095
%TLEAVE3	0.833531	T_CMPTR2	0.932962	EMP_CMI3	0.866534
%TLEAVE4	0.926846	T_CMPTR3	0.951694	EMP_CMI?	0.989481
%TLEAVE5	0.936514	T_CMPTR4	0.929141	EMP_MAP2	0.782270
%TLEAVE?	0.914397	T_CMPTR?	0.983154	EMP_MAP3	0.843428
RESOURC2	0.918109	T_PRBSL2	0.920219	EMP_MAP?	0.987216
RESOURC3	0.926323	T_PRBSL3	0.959644	CALUNR-N	0.911734
RESOURC4	0.973442	T_PRBSL4	0.945270	CALUNR-?	0.989448
RESOURC?	0.961745	T_PRBSL?	0.994473	CALTST-N	0.925377
TMCH-PAR	0.926352	T_REPPJ2	0.918680	CALTST-?	0.993649
TMCH-COM	0.989332	T_REPPJ3	0.950547	INSERVM2	0.900741
T_MATCR2	0.875050	T_REPPJ4	0.968524	INSERVM3	0.920667
T_MATCR3	0.796021	T_DISMA2	0.837434	INSERVM4	0.940300
T_MATTR2	0.860467	T_DISMA3	0.857713	INSERVM5	0.926087
T_MATTR3	0.860176	T_DISMA4	0.812257	INSERVM?	0.962030
AB_MATH2	0.935402	T_DISMA?	0.996229		
AB_MATH3	0.940844	T_RLPRB2	0.859841		
AB_MATH4	0.940498	T_RLPRB3	0.886254		
AB_MATH?	0.987528	T_RLPRB4	0.881907		
T_MATHW2	0.939714	T_MUPRB2	0.872030		
T_MATHW3	0.954898	T_MUPRB3	0.955449		

Table F-53
Proportion of Variance of the Conditioning Variable Contrasts Accounted for by the Principal Components Used in the Conditioning Model for Mathematics Main Conditioning Variables, Age 13/Grade 8

Contrast	Proportion of Variance	Contrast	Proportion of Variance	Contrast	Proportion of Variance
FEMALE	0.949744	CMP4SCH5	0.798135	S_REPPJ?	0.833934
BLACK	0.808765	ST8GRD12	0.939588	HVCALC-N	0.948787
HISPANIC	0.801460	ST8GRD35	0.946095	SCICAL-N	0.973219
ASIAN	0.945046	ST8GRD>5	0.880376	8GR_MAT8	0.907492
HI_METRO	0.887324	SLP_MA-L	0.750371	PRE-ALG8	0.961398
STOC-OTH	0.874353	S_TXTBK2	0.926785	ALGEBRA8	0.795871
S_EAST	0.863136	S_TXTBK3	0.949002	OTHER_M8	0.874159
CENTRAL	0.872660	S_TXTBK4	0.908251	BASIC9	0.983027
WEST	0.874161	S_TXTBK?	0.837933	PRE-ALG9	0.963052
HS_GRAD	0.967678	S_WRKSH2	0.904671	ALGEBRA9	0.962823
POST_HS	0.974711	S_WRKSH3	0.917210	GEOMTRY9	0.880421
COL_GRAD	0.960789	S_WRKSH4	0.917059	OTHER_M9	0.966925
PARED-?	0.979883	S_WRKSH?	0.869860	MCLASS9?	0.979491
HITEM=3	0.871550	S_SMGRP2	0.939247	S_MATHW1	0.930433
HITEM=4	0.841695	S_SMGRP3	0.943275	S_MATHW2	0.955223
TV-LIN	0.982545	S_SMGRP4	0.926645	S_MATHW3	0.991478
TV-QUAD	0.982922	S_SMGRP?	0.773576	S_MATHW4	0.980745
HL-SM/AL	0.852564	S_MI&GS2	0.970391	S_MATHW5	0.941530
HWC-NONE	0.958295	S_MI&GS3	0.962706	S_MATHW6	0.915457
HWC-YES	0.962950	S_MI&GS4	0.954262	S_MATHW?	0.614674
HW-LIN	0.937959	S_MI&GS?	0.785607	LKMAT-A	0.891396
HW-QUAD	0.916814	S_CALCR2	0.937356	LKMAT-U	0.929521
MINORITY	0.845337	S_CALCR3	0.932958	LKMAT-D	0.954690
INTEGRAT	0.858942	S_CALCR4	0.882302	LKMAT-SD	0.792656
=MA/<MG	0.878014	S_CALCR?	0.750979	GDMAT-A	0.880891
=MA/=MG	0.959743	S_CMPTR2	0.957710	GDMAT-U	0.932365
=MA/>MG	0.966659	S_CMPTR3	0.958961	GDMAT-D	0.962862
>MA/=MG	0.961473	S_CMPTR4	0.939252	GDMAT-SD	0.753756
NON_PUBL	0.817539	S_CMPTR?	0.656466	NSMAT-A	0.976617
2PARENTS	0.706684	S_PRBSL2	0.963212	NSMAT-U	0.967437
MOM@HM-Y	0.835796	S_PRBSL3	0.957411	NSMAT-D	0.964552
>=6_PGS	0.787509	S_PRBSL4	0.939128	NSMAT-SD	0.943642
>=11_PGS	0.794765	S_PRBSL?	0.763291	USMAT-A	0.859004
<=2_DAYS	0.969483	S_MUPRB2	0.978015	USMAT-U	0.902709
USA-NO/?	0.960837	S_MUPRB3	0.969252	USMAT-D	0.922535
CHGSCH=1	0.881595	S_MUPRB4	0.960408	USMAT-SD	0.675079
CHGSCH=2	0.909880	S_MUPRB?	0.718340	M4BOY-A	0.991437
CHGSCH3+	0.809370	S_MATST2	0.937761	M4BOY-U	0.987652
DIS@HOM2	0.875617	S_MATST3	0.941508	M4BOY-D	0.985319
DIS@HOM3	0.921179	S_MATST4	0.982820	M4BOY-SD	0.978834
DIS@HOM4	0.814755	S_MATST?	0.709494	MATMF-A	0.931633
CMP4SCH2	0.957074	S_REPPJ2	0.983385	MATMF-U	0.940335
CMP4SCH3	0.946030	S_REPPJ3	0.973825	MATMF-D	0.951188
CMP4SCH4	0.946090	S_REPPJ4	0.970835	MATMF-SD	0.753004

Table F-53 (continued)
Proportion of Variance of the Conditioning Variable Contrasts Accounted for by the Principal Components Used in the Conditioning Model for Mathematics Main Conditioning Variables, Age 13/Grade 8

Contrast	Proportion of Variance	Contrast	Proportion of Variance	Contrast	Proportion of Variance
MAJOB-A	0.904192	%REMRED4	0.862658	T_MATHW6	0.988299
MAJOB-U	0.943465	%REMRED5	0.896869	T_MATHW?	0.974204
MAJOB-D	0.961356	%REMRED6	0.826230	T_TXTBK2	0.802932
MAJOB-SD	0.806332	%REMRED7	0.928166	T_TXTBK3	0.875105
MAPRB-A	0.896458	%REMRED8	0.918110	T_TXTBK4	0.872540
MAPRB-U	0.922776	%REMRED?	0.912127	T_TXTBK?	0.988606
MAPRB-D	0.951199	%REMMAT2	0.819687	T_WRKSH2	0.881460
MAPRB-SD	0.786498	%REMMAT3	0.840086	T_WRKSH3	0.925370
#QUEST+2	0.922725	%REMMAT4	0.819905	T_WRKSH4	0.861865
#QUEST+3	0.909042	%REMMAT5	0.863630	T_WRKSH?	0.983572
#QUEST+4	0.940441	%REMMAT6	0.879377	T_SMGRP2	0.822370
TESTDIF2	0.945924	%REMMAT?	0.952052	T_SMGRP3	0.929999
TESTDIF3	0.943553	%ENR/YR2	0.866837	T_SMGRP4	0.817872
TESTDIF4	0.946497	%ENR/YR3	0.888681	T_SMGRP?	0.982451
TESTEFF2	0.931051	%ENR/YR4	0.821688	T_MI&GS2	0.920491
TESTEFF3	0.913544	%ENR/YR?	0.902116	T_MI&GS3	0.922570
TESTEFF4	0.939815	%RETAIN2	0.810001	T_MI&GS4	0.924142
TESTIMP2	0.909261	%RETAIN3	0.843841	T_MI&GS?	0.979619
TESTIMP3	0.892770	%RETAIN4	0.859978	T_CALCR2	0.894747
TESTIMP4	0.912851	%RETAIN5	0.840906	T_CALCR3	0.903847
TESTIMP?	0.976480	%RETAIN?	0.866273	T_CALCR4	0.889538
DSOLUTN2	0.927399	%TLEAVE2	0.900523	T_CALCR?	0.985848
DSOLUTN3	0.924684	%TLEAVE3	0.852380	T_CMPTR2	0.909855
DSOLUTN4	0.916387	%TLEAVE4	0.912773	T_CMPTR3	0.965806
ALG4HS-N	0.824248	%TLEAVE5	0.921901	T_CMPTR4	0.965806
ALG4HS-?	0.888304	%TLEAVE?	0.921418	T_CMPTR?	0.983802
ENG-T=1S	0.887793	RESOURC2	0.896142	T_PRBSL2	0.877643
MAT-T=1S	0.885416	RESOURC3	0.915351	T_PRBSL3	0.942490
RPRIOR-N	0.784482	RESOURC4	0.898339	T_PRBSL4	0.884708
RPRIOR-?	0.917306	RESOURC?	0.958557	T_PRBSL?	0.984350
WPRIOR-N	0.795177	TMCH-PAR	0.872867	T_REPPJ2	0.960098
WPRIOR-?	0.871399	TMCH-COM	0.981520	T_REPPJ3	0.942404
MPRIOR-N	0.792713	T_MATCR2	0.854590	T_REPPJ4	0.969840
MPRIOR-?	0.882067	T_MATCR3	0.856945	T_DISMA2	0.799559
%SUBLUN2	0.880356	T_MATTR2	0.876813	T_DISMA3	0.828707
%SUBLUN3	0.907545	T_MATTR3	0.913674	T_DISMA4	0.830690
%SUBLUN4	0.914389	AB_MATH2	0.912870	T_DISMA?	0.978522
%SUBLUN5	0.921319	AB_MATH3	0.915249	T_RLPRB2	0.875820
%SUBLUN6	0.905725	AB_MATH4	0.917759	T_RLPRB3	0.904292
%SUBLUN7	0.834618	AB_MATH?	0.973429	T_RLPRB4	0.886504
%SUBLUN8	0.876710	T_MATHW2	0.922098	T_RLPRB?	0.983184
%SUBLUN?	0.867480	T_MATHW3	0.964971	T_MUPRB2	0.867781
%REMRED2	0.836947	T_MATHW4	0.941682	T_MUPRB3	0.934424
%REMRED3	0.844724	T_MATHW5	0.927540	T_MUPRB4	0.918162

Table F-53 (continued)
Proportion of Variance of the Conditioning Variable Contrasts Accounted for by the Principal Components Used in the Conditioning Model for Mathematics Main Conditioning Variables, Age 13/Grade 8

Contrast	Proportion of Variance	Contrast	Proportion of Variance	Contrast	Proportion of Variance
T_MUPRB?	0.984835				
EMP_N&O2	0.773188				
EMP_N&O3	0.874350				
EMP_N&O?	0.983034				
EMP_MEA2	0.854184				
EMP_MEA3	0.889331				
EMP_MEA?	0.984287				
EMP_GEO2	0.829863				
EMP_GEO3	0.823146				
EMP_GEO?	0.979516				
EMP_DSP2	0.898354				
EMP_DSP3	0.920219				
EMP_DSP?	0.973807				
EMP_ALG2	0.797833				
EMP_ALG3	0.874897				
EMP_ALG?	0.974572				
EMP_F/C2	0.841993				
EMP_F/C3	0.869266				
EMP_F/C?	0.976801				
EMP_S/P2	0.801515				
EMP_S/P3	0.827957				
EMP_S/P?	0.984437				
EMP_R/A2	0.798159				
EMP_R/A3	0.912190				
EMP_R/A?	0.984720				
EMP_CMI2	0.799678				
EMP_CMI3	0.849740				
EMP_CMI?	0.971496				
EMP_MAP2	0.808713				
EMP_MAP3	0.818866				
EMP_MAP?	0.982819				
CALUNR-N	0.856461				
CALUNR-?	0.984565				
CALTST-N	0.841583				
CALTST-?	0.976082				
INSERVM2	0.923431				
INSERVM3	0.928873				
INSERVM4	0.912549				
INSERVM5	0.930035				
INSERVM?	0.948002				

Table F-54
Proportion of Variance of the Conditioning Variable Contrasts Accounted for by the Principal Components Used in the Conditioning Model for Mathematics Main Conditioning Variables, Age 17/Grade 12

Contrast	Proportion of Variance	Contrast	Proportion of Variance	Contrast	Proportion of Variance
FEMALE	0.807563	#LAN-LIN	0.710468	S_PRBSL3	0.962750
BLACK	0.752152	#VOC-Y	0.901448	S_PRBSL4	0.944812
HISPANIC	0.821220	#VOC-LIN	0.703839	S_PRBSL?	0.813671
ASIAN	0.812452	#ART-Y	0.897010	S_MUPRB2	0.983251
HI_METRO	0.896202	#ART-LIN	0.666570	S_MUPRB3	0.974255
STOC-OTH	0.899937	USA-NO/?	0.689350	S_MUPRB4	0.968155
S_EAST	0.836674	DIS@HOM2	0.894934	S_MUPRB?	0.764873
CENTRAL	0.890699	DIS@HOM3	0.913862	S_MATST2	0.936548
WEST	0.877390	DIS@HOM4	0.868514	S_MATST3	0.948629
HS_GRAD	0.967362	CMP4SCH2	0.933111	S_MATST4	0.842481
POST_HS	0.967385	CMP4SCH3	0.916518	S_MATST?	0.833742
COL_GRAD	0.958308	CMP4SCH4	0.907242	S_REPPJ2	0.991002
PARED-?	0.982480	CMP4SCH5	0.758460	S_REPPJ3	0.979940
HITEM=3	0.888180	VOC/BUSI	0.976323	S_REPPJ4	0.977868
HITEM=4	0.865275	2-YR_COL	0.963102	S_REPPJ?	0.923266
TV-LIN	0.977429	4-YR_COL	0.934664	HVCALC-N	0.954500
TV-QUAD	0.977710	MILITARY	0.968979	SCICAL-N	0.916043
HL-SM/AL	0.769555	OTHERACT	0.854817	MATH12-N	0.806752
HWC-NONE	0.966003	SLP_MA-L	0.786123	ALG1-<9	0.896465
HWC-YES	0.970818	S_TXTBK2	0.978064	ALG1-9	0.963216
HW-LIN	0.907464	S_TXTBK3	0.979939	ALG1-10	0.941428
HW-QUAD	0.873340	S_TXTBK4	0.810850	ALG1->10	0.919126
MINORITY	0.840988	S_TXTBK?	0.892269	ALG1-?	0.800892
INTEGRAT	0.904892	S_WRKSH2	0.909839	S_MATHW1	0.962646
=MA/<MG	0.934636	S_WRKSH3	0.939538	S_MATHW2	0.941035
=MA/=MG	0.967599	S_WRKSH4	0.887665	S_MATHW3	0.967959
>MA/=MG	0.983569	S_WRKSH?	0.930772	S_MATHW4	0.972265
NON_PUBL	0.826277	S_SMGRP2	0.937496	S_MATHW5	0.944961
2PARENTS	0.829874	S_SMGRP3	0.940557	S_MATHW6	0.927619
MOM@HM-Y	0.803534	S_SMGRP4	0.911494	S_MATHW?	0.785571
>=6_PGS	0.807137	S_SMGRP?	0.859147	LKMAT-A	0.863792
>=11_PGS	0.809177	S_MI&GS2	0.956644	LKMAT-U	0.934065
<=2_DAYS	0.972854	S_MI&GS3	0.952548	LKMAT-D	0.955823
COL_PREP	0.616723	S_MI&GS4	0.937059	LKMAT-SD	0.771166
VOC/TECH	0.898206	S_MI&GS?	0.854696	GDMAT-A	0.894535
#ENG-Y	0.860896	S_CALCR2	0.912317	GDMAT-U	0.946804
#ENG-LIN	0.873145	S_CALCR3	0.918751	GDMAT-D	0.965967
#MATH-Y	0.899599	S_CALCR4	0.816002	GDMAT-SD	0.803550
#MAT-LIN	0.896307	S_CALCR?	0.874914	USMAT-A	0.886568
#SCI-Y	0.886888	S_CMPTR2	0.958624	USMAT-U	0.937562
#SCI-LIN	0.844460	S_CMPTR3	0.938956	USMAT-D	0.954526
#HIS-Y	0.890184	S_CMPTR4	0.914965	USMAT-SD	0.715086
#HIS-LIN	0.805332	S_CMPTR?	0.783454	M4BOY-A	0.994564
#FLANG-Y	0.892850	S_PRBSL2	0.969428	M4BOY-U	0.975111

Table F-54 (continued)
Proportion of Variance of the Conditioning Variable Contrasts Accounted for by the Principal Components Used in the Conditioning Model for Mathematics Main Conditioning Variables, Age 17/Grade 12

Contrast	Proportion of Variance	Contrast	Proportion of Variance	Contrast	Proportion of Variance
M4BOY-D	0.986417	CALCLUS2	0.970278	%REMRED6	0.874090
M4BOY-SD	0.977007	CALCLUS3	0.985786	%REMRED?	0.898049
MATMF-A	0.956812	CALCLUS4	0.971010	%REMMAT2	0.855237
MATMF-U	0.970580	COMPSCI2	0.966269	%REMMAT3	0.823266
MATMF-D	0.963262	COMPSCI3	0.975213	%REMMAT4	0.883157
MATMF-SD	0.945403	COMPSCI4	0.954746	%REMMAT5	0.906636
MAJOB-A	0.856355	#QUEST+2	0.952161	%REMMAT6	0.847378
MAJOB-U	0.905467	#QUEST+3	0.948214	%REMMAT8	0.937579
MAJOB-D	0.927770	#QUEST+4	0.947142	%REMMAT?	0.870788
MAJOB-SD	0.795374	TESTDIF2	0.959145	%ENR/YR2	0.890994
MAPRB-A	0.878278	TESTDIF3	0.947764	%ENR/YR3	0.877202
MAPRB-U	0.902170	TESTDIF4	0.937212	%ENR/YR4	0.829750
MAPRB-D	0.943337	TESTEFF2	0.976115	%ENR/YR?	0.857051
MAPRB-SD	0.789791	TESTEFF3	0.959403	%TLEAVE2	0.822060
GENMATH2	0.856225	TESTEFF4	0.958041	%TLEAVE3	0.890709
GENMATH3	0.963272	TESTIMP2	0.956551	%TLEAVE4	0.928914
GENMATH4	0.780410	TESTIMP3	0.954217	%TLEAVE5	0.897279
BUSMATH2	0.949213	TESTIMP4	0.948193	%TLEAVE?	0.940454
BUSMATH3	0.971794	TESTIMP?	0.983698	%2YRCOL2	0.889823
BUSMATH4	0.934882	DSOLUTN2	0.950940	%2YRCOL3	0.884392
PRE-ALG2	0.912868	DSOLUTN3	0.940109	%2YRCOL4	0.912034
PRE-ALG3	0.967616	DSOLUTN4	0.921057	%2YRCOL5	0.865033
PRE-ALG4	0.923970	TRISCH-N	0.853852	%2YRCOL?	0.872421
ALGBR1-2	0.853155	TRISCH-?	0.926754	%4YRCOL2	0.894677
ALGBR1-3	0.959843	PCASCH-N	0.811166	%4YRCOL3	0.931923
ALGBR1-4	0.848276	PCASCH-?	0.918277	%4YRCOL4	0.954209
GEOMTRY2	0.829916	P/SSCH-N	0.870936	%4YRCOL5	0.910056
GEOMTRY3	0.980074	P/SSCH-?	0.824386	%4YRCOL6	0.908344
GEOMTRY4	0.829222	CALSCH-N	0.809435	%4YRCOL?	0.864755
ALGBR2-2	0.894173	CALSCH-?	0.898425		
ALGBR2-3	0.949930	CMPSCH-N	0.782022		
ALGBR2-4	0.844726	CMPSCH-?	0.897860		
TRIGTRY2	0.974645	%SUBLUN2	0.914756		
TRIGTRY3	0.891259	%SUBLUN3	0.945038		
TRIGTRY4	0.915330	%SUBLUN4	0.938953		
PRE-CAL2	0.980918	%SUBLUN5	0.945801		
PRE-CAL3	0.919462	%SUBLUN6	0.873745		
PRE-CAL4	0.950727	%SUBLUN7	0.943238		
PROBSTA2	0.982895	%SUBLUN8	0.840560		
PROBSTA3	0.972435	%SUBLUN?	0.788189		
PROBSTA4	0.976610	%REMRED2	0.862532		
UNIFMAT2	0.963123	%REMRED3	0.841483		
UNIFMAT3	0.962210	%REMRED4	0.871526		
UNIFMAT4	0.944944	%REMRED5	0.881713		

Table F-55
Proportion of Variance of the Conditioning Variable Contrasts Accounted for by the Principal Components Used in the Conditioning Model for Writing Main Conditioning Variables, Age 9/Grade 4

Contrast	Proportion of Variance	Contrast	Proportion of Variance	Contrast	Proportion of Variance
FEMALE	0.974022	CMP4SCH2	0.955326	IORG_SOM	0.916925
BLACK	0.823991	CMP4SCH3	0.970984	IORG_NEV	0.936796
HISPANIC	0.843050	CMP4SCH4	0.942871	IORG_MIS	0.852824
ASIAN	0.957203	CMP4SCH5	0.989123	ICRE_SOM	0.966058
HI_METRO	0.908483	LKW__AG	0.895577	ICRE_NEV	0.979420
STOC-OTH	0.914843	LKW__UND	0.917224	ICRE_MIS	0.823740
S_EAST	0.912549	LKW__DAG	0.936587	ILEN_SOM	0.881659
CENTRAL	0.906208	LKW_SDAG	0.945765	ILEN_NEV	0.868608
WEST	0.915053	LKW_MISS	0.721693	ILEN_MIS	0.841928
HS_GRAD	0.987261	GDW__AG	0.886482	CSP__MON	0.918280
POST_HS	0.991533	GDW__UND	0.916384	CSP__NEV	0.812102
COL_GRAD	0.976836	GDW__DAG	0.947146	CSP_MISS	0.899775
PARED-?	0.976410	GDW_SDAG	0.964696	CRP__MON	0.874402
HITEM=3	0.845425	GDW_MISS	0.803840	CRP__NEV	0.821730
HITEM=4	0.800992	WST__WK	0.943492	CRP_MISS	0.894803
TV-LIN1	0.987531	WST__MON	0.941080	DOWEL_PG	0.828769
TV-QUAD	0.987921	WST__NEV	0.954768	DOWEL_NG	0.856982
HL-SM/AL	0.951822	WST_MISS	0.972929	DOWEL_BD	0.929628
HW4-NONE	0.992752	TLK_SOME	0.917075	TESTDIF2	0.951935
HW4-YES	0.993183	TLK__NEV	0.913368	TESTDIF3	0.925024
HW4-LIN1	0.987803	TLK_MISS	0.770839	TESTDIF4	0.923902
HW4QUAD1	0.986722	DFT_SOME	0.817901	TESTEFF2	0.888508
MINORITY	0.818864	DFT__NEV	0.808232	TESTEFF3	0.839781
INTEGRAT	0.865958	DFT_MISS	0.800646	TESTEFF4	0.911755
=MA/<MG	0.992443	CLT_SOME	0.866519	TESTIMP2	0.947063
=MA/=MG	0.995526	CLT__NEV	0.860091	TESTIMP3	0.962944
=MA/>MG	0.996095	CLT_MISS	0.727699	TESTIMP4	0.973678
>MA/=MG	0.996599	SPG__WK	0.869767	TESTIMP?	0.968532
NON_PUBL	0.899853	SPG__MON	0.905217	ESSPOR_M	0.925684
2PARENTS	0.977647	SPG__NEV	0.863796	ESSPOR_Y	0.917702
2PARENTS	0.972291	SPG_MISS	0.761847	ESSPOR_N	0.902393
MOM@HM-Y	0.987893	GRP__WK	0.928515	RPRIOR-N	0.913236
6-10PGS	0.821695	GRP__MON	0.934396	RPRIOR-?	0.758234
>10_PGS	0.806443	GRP__NEV	0.921859	WPRIOR-N	0.821455
PRESCH-Y	0.982408	GRP_MISS	0.834947	WPRIOR-?	0.800757
USA-NO/?	0.981977	LOG__WK	0.897318	%SUBLUN2	0.904814
CHGSCH=1	0.863322	LOG__MON	0.911516	%SUBLUN3	0.905180
CHGSCH=2	0.884875	LOG__NEV	0.851041	%SUBLUN4	0.925660
CHGSCH3+	0.810355	LOG_MISS	0.775626	%SUBLUN5	0.823899
ST4GRD12	0.871058	PORTF__N	0.978554	%SUBLUN?	0.817313
ST4GRD3+	0.869482	PORTF__?	0.975946	%REMRED2	0.908639
DIS@HOM2	0.869359	ISG_SOME	0.955990	%REMRED3	0.889747
DIS@HOM3	0.925661	ISG__NEV	0.971093	%REMRED4	0.918217
DIS@HOM4	0.830403	ISG_MISS	0.802535	%REMRED?	0.803904

Table F-55 (continued)
Proportion of Variance of the Conditioning Variable Contrasts Accounted for by the Principal Components Used in the Conditioning Model for Writing Main Conditioning Variables, Age 9/Grade 4

Contrast	Proportion of Variance	Contrast	Proportion of Variance	Contrast	Proportion of Variance
RD_POL-N	0.855471				
RD_POL-?	0.904298				
WR_POL-N	0.861525				
WR_POL-?	0.839368				
PARAID-O	0.901222				
PARAID-N	0.885003				
PARAID-?	0.869615				
CMPCLA-N	0.899038				
CMPCLA-?	0.789858				
CMPLAB-N	0.938747				
CMPLAB-?	0.786981				
CMPLAB-N	0.885447				
CMPLAB-?	0.783325				
%BIL_ED2	0.931569				
%BIL_ED3	0.929346				
%BIL_ED4	0.906417				
%BIL_ED5	0.905024				
%BIL_ED6	0.890523				
%BIL_ED7	0.987009				
%BIL_ED8	0.872434				
%BIL_ED?	0.950924				
%ESL_IN2	0.902855				
%ESL_IN3	0.903078				
%ESL_IN4	0.939298				
%ESL_IN5	0.918698				
%ESL_IN6	0.973260				
%ESL_IN7	0.874002				
%ESL_IN8	0.960031				
%ESL_IN?	0.925186				
PARVIS-O	0.871586				
PARVIS-N	0.959258				
REQ_HW-N	0.836732				
REQ_HW-?	0.870887				

Table F-56
Proportion of Variance of the Conditioning Variable Contrasts Accounted for by the Principal Components Used in the Conditioning Model for Writing Main Conditioning Variables, Age 13/Grade 8

Contrast	Proportion of Variance	Contrast	Proportion of Variance	Contrast	Proportion of Variance
FEMALE	0.975169	CMP4SCH4	0.939505	ANL_MISS	0.830081
BLACK	0.832034	CMP4SCH5	0.857445	PER__WK	0.974500
HISPANIC	0.794345	ST8GRD12	0.967840	PER_MON	0.969717
ASIAN	0.934402	ST8GRD35	0.942961	PER_NEV	0.963910
HI_METRO	0.898892	ST8GRD>5	0.887535	PER_MISS	0.816316
STOC-OTH	0.882295	LRN__AG	0.939637	NAR__WK	0.962707
S_EAST	0.896914	LRN_UND	0.943550	NAR_MON	0.953281
CENTRAL	0.887746	LRN_DAG	0.954449	NAR_NEV	0.956777
WEST	0.899687	LRN_SDAG	0.979920	NAR_MISS	0.858445
HS_GRAD	0.966661	LRN_MISS	0.901128	PLA_SOME	0.887982
POST_HS	0.975266	TIN__AG	0.903906	PLA_NEV	0.866573
COL_GRAD	0.960281	TIN_UND	0.917608	PLA_MISS	0.795242
PARED-?	0.979536	TIN_DAG	0.955863	OTL_SOME	0.882219
HITEM=3	0.876561	TIN_SDAG	0.966702	OTL_NEV	0.839132
HITEM=4	0.844206	TIN_MISS	0.914236	OTL_MISS	0.804903
TV-LIN1	0.984379	LKW__AG	0.922066	AUD_SOME	0.868460
TV-QUAD1	0.984292	LKW_UND	0.943898	AUD_NEV	0.829210
HL-SM/AL	0.919060	LKW_DAG	0.959066	AUD_MISS	0.722540
HWC-NONE	0.971229	LKW_SDAG	0.924434	TLK_SOME	0.862725
HWC-YES	0.971064	LKW_MISS	0.697848	TLK_NEV	0.842693
HW-LIN1	0.976216	GDW__AG	0.931238	TLK_MISS	0.797071
HW-QUAD1	0.974803	GDW_UND	0.951544	SOU_SOME	0.838179
MINORITY	0.838543	GDW_DAG	0.970138	SOU_NEV	0.804896
INTEGRAT	0.873323	GDW_SDAG	0.916892	SOU_MISS	0.753725
=MA/<MG	0.817754	GDW_MISS	0.891078	DFT_SOME	0.858062
=MA/=MG	0.943146	PAR__WK	0.915273	DFT_NEV	0.826527
=MA/>MG	0.992386	PAR_MON	0.907727	DFT_MISS	0.748181
>MA/=MG	0.950733	PAR_NEV	0.928139	CLT_SOME	0.889646
NON_PUBL	0.838266	PAR_MISS	0.757949	CLT_NEV	0.878883
2PARENTS	0.966067	PGS__WK	0.930233	CLT_MISS	0.804439
2PARENTS	0.978283	PGS_MON	0.950537	SPG__WK	0.875223
MOM@HM-Y	0.866157	PGS_NEV	0.914019	SPG_MON	0.891611
6-10PGS	0.822150	PGS_MISS	0.740315	SPG_NEV	0.884014
>10_PGS	0.789020	3PP__WK	0.961472	SPG_MISS	0.880445
<=2_DAYS	0.991691	3PP_MON	0.965669	GRP__WK	0.947370
USA-NO/?	0.966277	3PP_NEV	0.947635	GRP_MON	0.947220
CHGSCH=1	0.979170	3PP_MISS	0.798708	GRP_NEV	0.944322
CHGSCH=2	0.986327	REP__WK	0.965454	GRP_MISS	0.914511
CHGSCH3+	0.963137	REP_MON	0.961543	LOG__WK	0.931569
DIS@HOM2	0.884694	REP_NEV	0.964447	LOG_MON	0.934688
DIS@HOM3	0.916416	REP_MISS	0.832053	LOG_NEV	0.873685
DIS@HOM4	0.830110	ANL__WK	0.975618	LOG_MISS	0.887782
CMP4SCH2	0.964275	ANL_MON	0.973025	PORTF_N	0.962627
CMP4SCH3	0.956373	ANL_NEV	0.972117	PORTF_?	0.673344

Table F-56 (continued)
Proportion of Variance of the Conditioning Variable Contrasts Accounted for by the Principal Components Used in the Conditioning Model for Writing Main Conditioning Variables, Age 13/Grade 8

Contrast	Proportion of Variance	Contrast	Proportion of Variance	Contrast	Proportion of Variance
ISG_SOME	0.939423	%REMRED2	0.833527	T_HISP	0.936357
ISG__NEV	0.919652	%REMRED3	0.867907	T_ASIAN	0.898846
ISG_MISS	0.861498	%REMRED4	0.851625	T_AM.IND	0.944924
IORG_SOM	0.895896	%REMRED?	0.904554	T_RACE-?	0.977989
IORG_NEV	0.826570	ENG/AB-N	0.857834	T_MEXICN	0.914594
IORG_MIS	0.903036	ENG/AB-?	0.912853	T_PUERTO	0.932166
ICRE_SOM	0.927065	CMPECL-N	0.797873	T_CUBAN	0.924216
ICRE_NEV	0.917019	CMPECL-?	0.879282	T_HISP-?	0.755863
ICRE_MIS	0.909448	CMPELB-N	0.866477	T_YREXP2	0.860914
ILEN_SOM	0.846609	CMPELB-?	0.919416	T_YREXP3	0.917616
ILEN_NEV	0.856163	CMPEBR-N	0.834333	T_YREXP4	0.932996
ILEN_MIS	0.913417	CMPEBR-?	0.894388	T_YREXP5	0.922573
CSP__MON	0.888899	PARAID-O	0.884816	T_YREXP?	0.988087
CSP__NEV	0.807907	PARAID-N	0.904493	TCERT-RG	0.906936
CSP_MISS	0.898406	PARAID-?	0.959785	TCERT-HI	0.921512
CRP__MON	0.891932	CMP_INS2	0.901451	TCERT-?	0.961759
CRP__NEV	0.884651	CMP_INS3	0.919175	CERTG-N	0.821872
CRP_MISS	0.886881	CMP_INS4	0.925452	CERTG-NS	0.868338
DOWEL_PG	0.893907	CMP_INS5	0.913781	CERTG-?	0.834541
DOWEL_NG	0.890368	CMP_INS?	0.821997	BACHELRS	0.921045
DOWEL_BD	0.949526	%BIL_ED2	0.864884	MASTERS	0.931589
TESTDIF2	0.987514	%BIL_ED3	0.922024	SPECLIST	0.916401
TESTDIF3	0.972319	%BIL_ED4	0.833858	DOCTORAT	0.914818
TESTDIF4	0.971791	%BIL_ED5	0.877387	PROFESSL	0.890741
TESTEFF2	0.930567	%BIL_ED6	0.932283	DEGREE-?	0.978790
TESTEFF3	0.918172	%BIL_ED7	0.902118	UGR_ED-Y	0.812329
TESTEFF4	0.946024	%BIL_ED8	0.887107	UGR_EG-Y	0.793123
TESTIMP2	0.918584	%BIL_ED?	0.897637	GR_ED-Y	0.853372
TESTIMP3	0.894435	%ESL_IN2	0.882306	GR_EG-Y	0.857054
TESTIMP4	0.902033	%ESL_IN3	0.854469	GR_NO-Y	0.845435
TESTIMP?	0.966610	%ESL_IN4	0.833643	T_NCLAS2	0.897188
ESSPOR_M	0.952066	%ESL_IN5	0.909161	T_NCLAS3	0.903811
ESSPOR_Y	0.946122	%ESL_IN6	0.834262	T_NCLAS4	0.915991
ESSPOR_N	0.949508	%ESL_IN?	0.937492	T_NCLAS5	0.931038
ENG-T=1S	0.781447	PARVIS-O	0.822232	T_NCLAS?	0.914848
RPRIOR-N	0.832179	PARVIS-N	0.830200	CNTL_IM2	0.882142
RPRIOR-?	0.927059	PARVIS-?	0.948673	CNTL_IM3	0.929801
WPRIOR-N	0.848572	REQ_HW-N	0.823463	CNTL_IM4	0.903313
WPRIOR-?	0.869574	REQ_HW-?	0.978792	CNTL_IM5	0.876105
%SUBLUN2	0.865991	TMCH-PAR	0.841645	CNTL_IM?	0.979942
%SUBLUN3	0.892946	TMCH-COM	0.985759	CNTL_CC2	0.861324
%SUBLUN4	0.893063	T_FEMALE	0.794775	CNTL_CC3	0.906641
%SUBLUN5	0.878437	T_SEX-?	0.988443	CNTL_CC4	0.871179
%SUBLUN?	0.898020	T_BLACK	0.884291	CNTL_CC5	0.865549

Table E-56 (continued)
Proportion of Variance of the Conditioning Variable Contrasts Accounted for by the Principal Components Used in the Conditioning Model for Writing Main Conditioning Variables, Age 13/Grade 8

Contrast	Proportion of Variance	Contrast	Proportion of Variance	Contrast	Proportion of Variance
CNTL_CC?	0.982204	WBK__SUP	0.922575	ASPESS_?	0.961507
CNTL_CS2	0.873674	WBK___NO	0.922513	ASNE__WK	0.883121
CNTL_CS3	0.857681	WBK____?	0.984082	ASNE_MON	0.943465
CNTL_CS4	0.867466	TXT__SUP	0.856844	ASNE_NEV	0.908926
CNTL_CS5	0.868719	TXT___NO	0.863149	ASNESS_?	0.966899
CNTL_CS?	0.978800	TXT____?	0.989871	ASJO__WK	0.901678
JHSENG-N	0.808912	GRMI_SUP	0.815581	ASJO_MON	0.885792
JHSENG-?	0.872455	GRMI__NO	0.845489	ASJO_NEV	0.916038
CRTOTH-N	0.870255	GRMI___?	0.947478	ASJOUR_?	0.967026
CRTOTH-?	0.855142	PROC_SUP	0.852808	IMPGMR_M	0.820897
UG_EGE-Y	0.847371	PROC__NO	0.863718	IMPGMR_U	0.880825
UG_OTH-Y	0.845424	PROC___?	0.969890	IMPGMR_?	0.985900
GR_OTH-Y	0.850197	R+WI_SUP	0.849050	IMPORG_M	0.829056
WRT_TR-O	0.845862	R+WI__NO	0.875762	IMPORG_U	0.878780
WRT_CR-1	0.872466	R+WI___?	0.974583	IMPORG_?	0.990047
WRT_CR-2	0.931068	LITI_SUP	0.854183	IMPCRE_M	0.829963
WRT_CR-3	0.929954	LITI__NO	0.852845	IMPCRE_U	0.824111
WRT_CR-4	0.925187	LITI___?	0.969149	IMPCRE_?	0.985079
WRT_CR-?	0.974218	ACSB_SUP	0.903435	IMPLEN_M	0.934700
CONF__<6	0.859761	ACSB__NO	0.900063	IMPLEN_U	0.942205
CON_6-15	0.865967	ACSB___?	0.946045	IMPLEN_?	0.984807
CONF_16+	0.867965	TPAR__WK	0.834089	IMPPUR_M	0.833275
CONF___?	0.969907	TPAR_MON	0.830010	IMPPUR_U	0.847877
RDPB__<6	0.905926	TPAR_NEV	0.850301	IMPPUR_?	0.989938
RDP_6-15	0.940731	TPAR___?	0.914277	ASKPLN_S	0.854582
RDPB_16+	0.909655	TPPS__WK	0.856506	ASKPLN_N	0.806458
RDPB___?	0.986007	TPPS_MON	0.907856	ASKPLN_?	0.979509
ST_AB__N	0.810420	TPPS_NEV	0.858075	ASKOUT_S	0.901506
ST_AB__?	0.982749	TPPS___?	0.915902	ASKOUT_N	0.902586
ST_AB__A	0.907847	TPP3__WK	0.911541	ASKOUT_?	0.979522
ST_AB__L	0.919705	TPP3_MON	0.886364	ASKAUD_S	0.794335
ST_AB__M	0.915595	TPP3_NEV	0.903438	ASKAUD_N	0.843208
ST_AB__?	0.979868	TPPS___?	0.888964	ASKAUD_?	0.986398
GR_AB__N	0.755597	ASGR__WK	0.911757	ASKDSC_S	0.805377
GR_AB__?	0.981353	ASGR_MON	0.944182	ASKDSC_N	0.827176
TWOUT_<1	0.893314	ASGR_NEV	0.888470	ASKDSC_?	0.984864
TWOUT__1	0.928575	ASGREP_?	0.972317	ASKDSF_S	0.902870
TWOUT__2	0.939379	ASGE__WK	0.897772	ASKDSF_N	0.906533
TWOUT_3+	0.876786	ASGE_MON	0.933983	ASKDSF_?	0.981255
TWOUT__?	0.985277	ASGE_NEV	0.879999	ASKDFT_S	0.802038
HSTW__60	0.882295	ASGESS_?	0.966964	ASKDFT_N	0.890753
HSTW__90	0.906876	ASPE__WK	0.905434	ASKDFT_?	0.990621
HSTW_120	0.891834	ASPE_MON	0.953998	ASKCOM_S	0.831455
HSTW___?	0.987158	ASPE_NEV	0.915113	ASKCOM_N	0.837728

Table F-56 (continued)
Proportion of Variance of the Conditioning Variable Contrasts Accounted for by the Principal Components Used in the Conditioning Model for Writing Main Conditioning Variables, Age 13/Grade 8

Contrast	Proportion of Variance	Contrast	Proportion of Variance	Contrast	Proportion of Variance
ASKCOM_?	0.988683	WTCP_NEV	0.918971		
ASKRES_S	0.795625	WTCOMP_?	0.972947		
ASKRES_N	0.881230	AVCP_YDF	0.811162		
ASKRES_?	0.984010	AVCP_YCL	0.860701		
ASKTLK_S	0.771219	AVCOMP_?	0.967192		
ASKTLK_N	0.836261	TPORT__N	0.805227		
ASKTLK_?	0.990168	TPORT__?	0.981043		
ASKCHO_S	0.839908				
ASKCHO_N	0.867736				
ASKCTB_S	0.893246				
ASKCTB_N	0.897777				
ASKCTB_?	0.980531				
DOSP__WK	0.833351				
DOSP_MON	0.890804				
DOSP_NEV	0.862969				
DOSPEL_?	0.984722				
DOGP__WK	0.913183				
DOGP_MON	0.938641				
DOGP_NEV	0.870387				
DOGRPS_?	0.989925				
DOJR__WK	0.894620				
DOJR_MON	0.887654				
DOJR_NEV	0.916270				
DOJOUR_?	0.990250				
ASMC__WK	0.870142				
ASMC_MON	0.919655				
ASMC_NEV	0.932544				
ASSEMC_?	0.953567				
ASSH__WK	0.891732				
ASSH_MON	0.944472				
ASSH_NEV	0.885697				
ASSESH_?	0.967203				
ASES__WK	0.873336				
ASES_MON	0.925354				
ASES_NEV	0.906015				
ASSEES_?	0.959832				
ASPF__WK	0.902020				
ASPF_MON	0.925503				
ASPF_NEV	0.883992				
ASSEPF_?	0.965574				
EXCP_MON	0.852358				
EXCP_NEV	0.898602				
EXCOMP_?	0.972211				
WTCP_MON	0.894377				

Table F-57

Proportion of Variance of the Conditioning Variable Contrasts Accounted for by the Principal Components Used in the Conditioning Model for Writing Main Conditioning Variables, Age 17/Grade 12

Contrast	Proportion of Variance	Contrast	Proportion of Variance	Contrast	Proportion of Variance
FEMALE	0.976303	CMP4SCH4	0.893593	ANL__MON	0.962401
BLACK	0.823305	CMP4SCH5	0.727446	ANL__NEV	0.965398
HISPANIC	0.804545	VOC/BUSI	0.975069	ANL_MISS	0.896582
ASIAN	0.819409	2-YR_COL	0.966388	PER___WK	0.984846
HI_METRO	0.869478	4-YR_COL	0.936314	PER__MON	0.982315
STOC-OTH	0.883118	MILITARY	0.984092	PER__NEV	0.980607
S_EAST	0.870659	OTHERACT	0.964331	PER_MISS	0.828433
CENTRAL	0.906188	LRN___AG	0.936921	NAR___WK	0.976097
WEST	0.874429	LRN__UND	0.950261	NAR__MON	0.968870
HS_GRAD	0.965273	LRN__DAG	0.944741	NAR__NEV	0.966642
POST_HS	0.966712	LRN_SDAG	0.977319	NAR_MISS	0.879116
COL_GRAD	0.955599	LRN_MISS	0.906320	PLA_SOME	0.885149
PARED-?	0.980041	TIN___AG	0.923997	PLA__NEV	0.801785
HITEM=3	0.883201	TIN__UND	0.933083	PLA_MISS	0.884345
HITEM=4	0.859006	TIN__DAG	0.961962	OTL_SOME	0.887612
TV-LIN1	0.980276	TIN_SDAG	0.980557	OTL__NEV	0.832551
TV-QUAD1	0.980461	TIN_MISS	0.910637	OTL_MISS	0.877117
HL-SM/AL	0.792242	LKW___AG	0.887902	AUD_SOME	0.858223
HWC-NONE	0.983741	LKW__UND	0.925731	AUD__NEV	0.785520
HWC-YES	0.983998	LKW__DAG	0.950335	AUD_MISS	0.788593
HW-LIN1	0.970800	LKW_SDAG	0.817528	TLK_SOME	0.856109
HW-QUAD1	0.968041	LKW_MISS	0.758777	TLK__NEV	0.804068
MINORITY	0.829777	GDW___AG	0.923779	TLK_MISS	0.858426
INTEGRAT	0.930916	GDW__UND	0.946507	SOU_SOME	0.863483
=MA/<MG	0.990213	GDW__DAG	0.971322	SOU__NEV	0.784482
=MA/=MG	0.991196	GDW_SDAG	0.775281	SOU_MISS	0.840345
>MA/=MG	0.992984	GDW_MISS	0.883413	DFT_SOME	0.844475
NON_PUBL	0.812253	PAR___WK	0.914842	DFT__NEV	0.781016
2PARENTS	0.974826	PAR__MON	0.912838	DFT_MISS	0.814955
2PARENTS	0.975490	PAR__NEV	0.930073	CLT_SOME	0.916018
MOM@HM-Y	0.942117	PAR_MISS	0.803128	CLT__NEV	0.910150
6-10PGS	0.845851	PGS___WK	0.949153	CLT_MISS	0.846058
>10_PGS	0.786539	PGS__MON	0.962492	SPG___WK	0.933111
<=2_DAYS	0.990900	PGS__NEV	0.930504	SPG__MON	0.942943
COL_PREP	0.628679	PGS_MISS	0.711201	SPG__NEV	0.920742
VOC/TECH	0.943898	3PP___WK	0.970662	SPG_MISS	0.893607
#ENG-Y	0.890468	3PP__MON	0.979657	GRP___WK	0.955836
#ENG-L01	0.868290	3PP__NEV	0.945465	GRP__MON	0.957288
USA-NO/?	0.699051	3PP_MISS	0.776212	GRP__NEV	0.957278
DIS@HOM2	0.894819	REP___WK	0.950777	GRP_MISS	0.910513
DIS@HOM3	0.912792	REP__MON	0.948441	LOG___WK	0.918100
DIS@HOM4	0.873363	REP__NEV	0.961106	LOG__MON	0.928503
CMP4SCH2	0.930719	REP_MISS	0.903613	LOG__NEV	0.874021
CMP4SCH3	0.911110	ANL___WK	0.959843	LOG_MISS	0.904867

Table F-57 (continued)
Proportion of Variance of the Conditioning Variable Contrasts Accounted for by the Principal Components Used in the Conditioning Model for Writing Main Conditioning Variables, Age 17/Grade 12

Contrast	Proportion of Variance	Contrast	Proportion of Variance	Contrast	Proportion of Variance
PORTF__N	0.975436	%REMRED?	0.930147		
PORTF__?	0.693956	ENG/AB-N	0.861646		
ISG_SOME	0.951946	ENG/AB-?	0.973839		
ISG__NEV	0.903695	#SEMENG8	0.750497		
ISG_MISS	0.880894	CMPECL-N	0.853431		
IORG_SOM	0.917981	CMPECL-?	0.887588		
IORG_NEV	0.807143	CMPELB-N	0.919836		
IORG_MIS	0.907980	CMPELB-?	0.936757		
ICRE_SOM	0.971129	CMPEBR-N	0.841273		
ICRE_NEV	0.931305	CMPEBR-?	0.858613		
ICRE_MIS	0.925335	AP_ENGL2	0.945895		
ILEN_SOM	0.851736	AP_ENGL3	0.936874		
ILEN_NEV	0.849223	AP_ENGL4	0.885565		
ILEN_MIS	0.922457	AP_ENGL5	0.915353		
CSP__MON	0.883777	AP_ENGL6	0.894162		
CSP__NEV	0.862727	AP_ENGL?	0.921755		
CSP_MISS	0.890604	PARAID-O	0.942532		
CRP__MON	0.913583	PARAID-N	0.941772		
CRP__NEV	0.891955	%BIL_ED2	0.934339		
CRP_MISS	0.876030	%BIL_ED3	0.945034		
DOWEL_PG	0.922117	%BIL_ED4	0.879942		
DOWEL_NG	0.929221	%BIL_ED5	0.850708		
DOWEL_BD	0.954652	%BIL_ED6	0.790726		
TESTDIF2	0.993447	%BIL_ED7	0.900620		
TESTDIF3	0.980360	%BIL_ED8	0.985773		
TESTDIF4	0.978613	%BIL_ED?	0.916461		
TESTEFF2	0.982504	%ESL_IN2	0.853923		
TESTEFF3	0.971496	%ESL_IN3	0.885563		
TESTEFF4	0.974649	%ESL_IN4	0.820684		
TESTIMP2	0.954950	%ESL_IN5	0.803276		
TESTIMP3	0.941056	%ESL_IN6	0.798033		
TESTIMP4	0.945889	%ESL_IN?	0.954107		
TESTIMP?	0.825643	PARVIS-O	0.809153		
ESSPOR_M	0.936436	PARVIS-N	0.821945		
ESSPOR_Y	0.936074	PARVIS-?	0.961027		
ESSPOR_N	0.938106	REQ_HW-N	0.789811		
%SUBLUN2	0.889204				
%SUBLUN3	0.909616				
%SUBLUN4	0.874883				
%SUBLUN5	0.882435				
%SUBLUN?	0.794259				
%REMRED2	0.869261				
%REMRED3	0.866116				
%REMRED4	0.844482				

APPENDIX G

THE NAEP ACHIEVEMENT LEVEL-SETTING PROCESS FOR THE 1992 MATHEMATICS ASSESSMENT

APPENDIX G

The NAEP Achievement Level-setting Process for the 1992 Mathematics Assessment[1]

Mary Lyn Bourque

National Assessment Governing Board

Introduction

Since 1984 NAEP has reported the performance of students in the nation and for specific subpopulations on a 0-to-500 proficiency scale. The history and development of the scale and the anchoring procedure used to interpret specific points on that scale is described elsewhere in this report.

However, the 1988 legislation[2] created an independent board, the National Assessment Governing Board (NAGB), responsible for setting policy for the NAEP program. The Board has a statutory mandate to identify "appropriate achievement goals for each. . .grade in each subject area to be tested under the National Assessment." Consistent with this directive, and striving to achieve one of the primary mandates of the statute "to improve the form and use of NAEP results," the Board set performance standards (called achievement levels by NAGB) for the National Assessment in 1990 and again in 1992.

The 1990 trial, initiated in December 1989 with the dissemination of a draft policy statement (NAGB, 1989) and culminating 22 months later in the publication of the NAGB report, *The Levels of Mathematics Achievement* (Bourque & Garrison, 1991), consisted of two phases: the main study and a replication-validation study. Although there were slight differences between the two phases, there were many common elements. Both phases used a modified (iterative/empirical) Angoff (1971) procedure for arriving at the levels; both focused on estimating performance levels based on a review of the 1990 NAEP mathematics item pool; and both phases employed a set of policy definitions for Basic, Proficient, and Advanced (NAGB,

---

[1] Several studies have examined the validity of the achievement level-setting process, including: 1) American College Testing. (1992). *Setting achievement levels on the 1992 National Assessment of Education Progress in mathematics, reading, and writing: A technical report on the reliability and validity.* Iowa City, IA: Author. 2) Burstein, L., Koretz, D. M., Linn, R. L., Sugrue, B., Novak, J., Lewis, E., & Baker, E. L. (1993). *The validity of interpretations of the 1992 NAEP achievement levels in mathematics.* Los Angeles, CA: CRESST. 3) National Academy of Education. (1993). *Setting performance standards for student achievement.* Stanford, CA: Author. 4) National Academy of Education. (1994). *Setting performance standards for student achievement: Background studies.* Stanford, CA: Author. 5) National Academy of Education. (1993). *The Trial State Assessment: Prospects and realities.* Stanford, CA: Author.

[2] Public Law 100-297. (1988). National assessment of educational progress improvement act (Article No. USC 1221). Washington, DC.

1990) as the criteria for making the item ratings. However, the 1990 process was evaluated by a number of different groups (see Hambleton & Bourque, 1991) who identified technical flaws in the 1990 process. These evaluations influenced NAGB's decision to set the levels again in 1992 and to not use the 1990 levels as benchmarks for progress toward the national goals during the coming decade. However, it is interesting to note that the 1990 and 1992 processes produced remarkably similar results.

In September 1991 NAGB contracted with American College Testing (ACT) to convene the panels of judges that would recommend the levels on the 1992 NAEP assessments in reading, writing, and mathematics. While the 1992 level-setting activities were not unlike those undertaken by NAGB in 1990, there were significant improvements made in the process for 1992. There was a concerted effort to bring greater technical expertise to the process: The contractor selected by NAGB has a national reputation for setting standards in a large number of certification and licensure exams; an internal and external advisory team monitored all the technical decisions made by the contractor throughout the process; and state assessment directors periodically provided their expertise and technical assistance at key stages in the project.

Setting achievement levels is a method for setting standards on the NAEP assessment that identifies what students should know and be able to do at various points along the proficiency scale. The initial policy definitions of the achievement levels were presented to panelists along with an illustrative framework for more in-depth development and operationalization of the levels. Panelists were asked to determine descriptions/definitions of the three levels from the specific framework developed for the NAEP assessment with respect to the content and skills to be assessed. The operationalized definitions were refined throughout the level-setting process, as well as validated with a supplementary group of judges subsequent to the level-setting meetings. Panelists were also asked to develop a list of illustrative tasks associated with each of the levels, after which sample items from the NAEP item pool were identified to exemplify the full range of performance of the intervals between levels. The emphasis in operationalizing the definitions and in identifying and selecting exemplar items and papers was to represent the full range of performance from the lower level to the next higher level. The details of the implementation procedures are outlined in the remainder of this appendix.

Preparing for the 1992 Mathematics Level-setting Meeting

It is important for the planning of any standard-setting effort to know how various process elements interact with each other. For example, panelists interact with pre-meeting materials, the meeting materials (i.e., the assessment questions, rating forms, rater feedback, and so forth), each other, and the project staff. All of these elements combine to promote or degrade what has been called intrajudge consistency and interjudge consensus (Friedman & Ho, 1990).

Previous research has conceptualized the effects of two major kinds of interaction: people interacting with text (Smith & Smith, 1988) and people interacting with each other (Curry, 1987; Fitzpatrick, 1989). To assess the effects of textual and social interaction and adjust the standard setting procedures accordingly, a pilot study was conducted as the first phase of the 1992 initiative.

Reading was chosen as the single content area to be pilot-tested since it combined all of the various features found in the other NAEP assessments, including multiple-choice, short constructed-response, and extended constructed-response items. The pilot study provided the opportunity to implement and evaluate all aspects of the operational plan—background materials, meeting materials, study design, meeting logistics, staff function, and participant function.

The overall pilot effort was quite successful. The level-setting process worked well, and the pilot allowed the contractor to make improvements in the design before implementation activities began. For example, schedule changes were made that allowed the panelists more time to operationalize the policy definitions before beginning the item-rating task. Also, the feedback mechanisms used to inform panelists about interjudge and intrajudge consistency data were improved for clarity and utility to the entire process.

The Mathematics Level-setting Panel

Sixty-nine panelists representing 32 jurisdictions (31 states and the District of Columbia) from the 424 nominees were invited to participate in the level-setting process. They represented mathematics teachers at grades 4, 8, and 12, nonteacher educators, and members of the noneducator (general public) community. The group was balanced by gender, race/ethnicity, NAEP regions of the country, community type (low SES, not low SES), district size, and school type (public/private). One panelist was unable to attend due to a family emergency, resulting in 68 participants: 24 at grade 4 and 22 at grades 8 and 12.

Process for Developing the Achievement Levels

The four-and-one-half day session began with a brief overview of NAEP and NAGB, a presentation on the policy definitions of the achievement levels, a review of the NAEP mathematics assessment framework, and a discussion of factors that influence item difficulty. The purpose of the presentation was to focus panelists' attention on the mathematics framework and to emphasize the fact that panelists' work was directly related to the NAEP assessment, not to the whole domain of mathematics.

All panelists completed and self-scored an appropriate grade-level form of the NAEP assessment. The purpose of this exercise was to familiarize panelists with the test content and scoring protocols before beginning to develop the preliminary operationalized descriptions of the three levels.

Working in small groups of five or six, panelists expanded and operationalized the policy definitions of Basic, Proficient, and Advanced in terms of specific mathematical skills, knowledge, and behaviors that were judged to be appropriate expectations for students in each grade, and were in accordance with the current mathematics assessment framework.

The policy definitions are as follows:

Basic	This level, below proficient, denotes partial mastery of the knowledge and skills that are fundamental for proficient work at each grade—4, 8, and 12.
Proficient	This central level represents solid academic performance for each grade tested—4, 8, and 12. Students reaching this level have demonstrated competency over challenging subject matter and are well prepared for the next level of schooling.
Advanced	This higher level signifies superior performance beyond proficient grade-level mastery at grades 4, 8, and 12.

The small groups were allowed to brainstorm about what student performance *should* be, using the framework and their experience in completing the NAEP assessment as guides[3]. A comprehensive listing of grade-level descriptors was developed, and panelists were asked to identify the five that best described what students *should* be able to do at each of the levels. Those descriptors appearing with the greatest frequency were compiled into a discussion list for the grade-level groups. Additions, deletions, and modifications were made as a result of discussions, and the groups reached general agreement that the final list of descriptors represented what students *should* be able to do at each achievement level.

Panelists next received two hours of training in the Angoff method. Training was customized to reflect the unique item formats of the particular subject area assessment. Once a conceptual consensus was reached about the characteristics of *marginally* acceptable examinees at each of the three levels, practice items from the released pool were rated by the panelists according to the process defined in the contractor's plan. For multiple-choice and short constructed-response items, panelists were asked to rate each item for the expected probability of a correct response for a group of *marginally* acceptable examinees at the Basic, Proficient, and Advanced levels. For extended constructed-response items, panelists were asked to review 20 to 25 student response papers and select three papers, one for each achievement level, that typified *marginally* acceptable examinee performance for that level.

Following training in the Angoff method, the judges began the rating process, inspecting and rating each item in the pool for the expected probabilities of answering the item correctly at each level. Panelists completed three rounds of item ratings. For Round 1, panelists first answered the items in each section, then reviewed their answers using scoring keys and protocols. This process helped ensure that panelists would be thoroughly familiar with each item, including the foils and scoring rubrics, before rating the items. Panelists provided item ratings/paper selections for all three achievement levels, one item at a time, for all the items in a section, then proceeded to the next set of items, for which the process was repeated. During Round 1, panelists used their lists of descriptors and other training materials for guidance in the rating process.

---

[3] The panelists also reviewed about half the item pool (the half they would not be rating later) so the descriptors could be further modified if appropriate.

Following Round 1, item response theory (IRT) was used to convert the rating results[4] for each rater to a latent ability scale represented by the Greek letter theta (θ). This θ scale was the same scale used to calibrate the NAEP items evaluated by each panelist. In order to provide meaningful feedback about item ratings, a special *relative scale* was constructed, which was a linear transformation of the theta scale having a mean of 75 and standard deviation of 15. Before Round 2 of the rating process, panelists were given interjudge consistency information using this relative scale. This information allowed panelists to see on the scale where their individual mean item ratings were, relative to the mean for the group and to the means for other panelists. Reasons for extreme mean ratings, including the possibility that some panelists misinterpreted the item rating task, were discussed briefly.

Before Round 2, panelists were also given item difficulty data. This information was presented as the percentage of students who answered each item correctly during the actual NAEP administration, for items scored "correct" or "incorrect" (i.e., multiple-choice and short constructed-response items), and as the percentage of students receiving scores of 1, 2, 3, and 4 for the extended constructed-response items[5]. Panelists were told that this item difficulty information should be used as a reality check. For items on which item ratings differed substantially from the item difficulty value, panelists were asked to reexamine the item to determine if they had misinterpreted the item or misjudged its difficulty. Results of the data analysis, and panelists' own evaluations, indicated that the item difficulty information was perceived as very useful but had little impact on panelists' ratings.

For Round 2, panelists reviewed the same set of items they had rated in Round 1 and, using the interjudge consistency information, the item difficulty information, and the information provided prior to Round 1, they either confirmed their initial item ratings or adjusted their ratings to reflect the additional information. About one-third of Round 1 item ratings were adjusted during Round 2.

Following Round 2, panelists' ratings were reanalyzed and additional information was presented to panelists concerning intrajudge variability prior to Round 3. For each panelist, the intrajudge variability information consisted of those items that they had rated differently than items having similar difficulty, taking into consideration the panelist's aggregated item ratings. That is, the panelists' aggregated item ratings were converted to the theta (θ) scale. All items rated by the panelists were then analyzed in terms of the panelist's achievement level (θ) in comparison to actual student performance on the items. The observed item rating from each panelist was contrasted to an expected item rating. Those items with large differences between observed and expected ratings were identified. Panelists were given this information and asked to review each of these items and decide if their Round 2 ratings still accurately reflected their best judgments of the items. The intrajudge consistency data was to be used to flag items for reconsideration in the final round of rating.

---

[4] Because the IRT item parameters were not available for the polytomously scored (extended constructed-response) items, these items (five at grade 4, six each at grades 8 and 12) were not included in the following discussion of results.

[5] The percentages presented to the raters summed to 100 percent, but this excluded the percentages—around 80 percent, in some cases—of students who wrote responses that were judged to be "off-task," those who "skipped" that question and continued beyond that question, and those who, apparently, "never reached" that question.

For Round 3, panelists reviewed the same set of items they rated in Rounds 1 and 2 using both the new intrajudge variability information and the information made available during Rounds 1 and 2. In addition, panelists could discuss, within their small groups, ratings of specific items about which they were unsure. About 20 percent of the item ratings were adjusted during Round 3.

Process of Selecting Exemplar Items

Following the standard-setting meeting, a series of procedures was implemented to select exemplar items. First, expected and empirical p-values were computed for each item in the released item pool. Expected p-values were based on predicted performance at the cut-off score for each achievement level and empirical p-values were based on the average performance of all students responding to the item. Items that did not have expected p-values ≥ 0.51 for any of the levels were deleted from the item pool. Second, items were compared to the operationalized descriptions of the levels. Items that did not match the content of the descriptions were deleted from the item pool. Third, the remaining items were classified as possible Basic, Proficient, or Advanced exemplars based on content match. Fourth, the validation panel reviewed the items and recommended a set of items to serve as exemplars for the levels. The final set of items was reviewed and approved by NAGB at their May 1992 meeting. These procedures are described in detail below.

Using the standard-setting ratings, expected p-values were computed for each item at the cut point for each achievement level. The criteria described below were applied to the scale-level results and an analysis was conducted to delineate items that could serve as exemplars for each achievement level (Basic, Proficient, Advanced). More specifically, for an item to be chosen as a possible exemplar for the Basic achievement level:

1) The expected p-value for students at the cut point for the Basic level of achievement had to be greater than 0.51;

2) The content of the item had to match the content of the operationalized description of Basic; and

3) The empirical p-value for the item had to be higher than empirical p-values for items selected as exemplars for the Proficient level.

As an example:

Grade 4 Basic Level Item M022801

Level	Basic	Proficient	Advanced
Scale point	211	248	280
Expected p-value	0.70	0.82	0.94
Empirical p-value = 0.52			

For an item to be chosen as a possible exemplar for the Proficient achievement level:

1) The expected p-value for students at the cut-off score for the Proficient level of achievement had to be greater than 0.51;

2) The content of the item had to match the content of the operationalized description of Proficient; and

3) The empirical p-value for the item had to be lower than empirical p-values for Basic exemplar items, but higher than student p-values for Advanced exemplar items.

As an example:

Grade 4 Proficient Level Item M022001			
Level	**Basic**	**Proficient**	**Advanced**
Scale point	211	248	280
Expected p-value	0.37	0.58	0.76
Empirical p-value = 0.35			

For an item to be chosen as a possible exemplar for the Advanced achievement level:

1) The expected p-value for students at the cut-point for the Advanced level of achievement had to be greater than 0.51;

2) The content of the item had to match the content of the operationalized description of Advanced; and

3) The empirical p-value for the item had to be lower than empirical p-values for Proficient exemplar items.

As an example:

Grade 4 Advanced Level Item M023101			
Level	**Basic**	**Proficient**	**Advanced**
Scale point	211	248	280
Expected p-value	0.29	0.43	0.61
Empirical p-value = 0.22			

The analysis procedures described above yielded 31 items as possible grade 4 exemplars, 43 items as possible grade 8 exemplars, and 37 items as possible grade 12 exemplars, as follows:

Possible Exemplar Items by Grade and Achievement Level			
Grade	**Basic**	**Proficient**	**Advanced**
4	9	14	8
8	23	15	5
12	14	16	7

For grade 4, the possible exemplars represented 49 percent of the released item pool. For grades 8 and 12, the possible exemplars represented 54 percent of the released item pool for each grade.

Process for Validating the Levels

Eighteen mathematics educators participated in the item selection and content validation process. Ten of the panelists were mathematics teachers who had participated in the original achievement levels-setting process and who had been identified as outstanding panelists by grade group facilitators during this meeting. The other eight panelists represented the National Council of Teachers of Mathematics, the Mathematical Sciences Education Board, and state-level mathematics curriculum supervisors. To the extent possible, the group was balanced by race/ethnicity, gender, community type, and region of the country.

The two-and-one-half day meeting began by briefing panelists on the purpose of the meeting. They first reviewed the operationalized descriptions of the achievement levels for consistency with the NAGB policy definitions of Basic, Proficient, and Advanced and with the NAEP *Mathematics Objectives*. Next, they reviewed the operationalized descriptions of the achievement levels for qualities such as within- and across-grade consistency, grade-level appropriateness, and utility for increasing the public's understanding of the NAEP mathematics results. Finally, working first in grade level (4, 8, and 12) groups of six panelists each, then as a whole group, panelists revised the operationalized descriptions to provide more within- and across-grade consistency and to align the language of the description more closely with the language of the NCTM *Standards*. Both the original descriptions and the revised descriptions are included later in this appendix.

On the third day, panelists again split into grade-level groups of six panelists each and reviewed the possible exemplar items. The task was to select a set of items, for each achievement level for their grade, that would best communicate to the public the levels of mathematics ability and the types of skills needed to perform in mathematics at that level.

After selecting sets of items for their grades, the three grade-level groups met as a whole group to review item selection. During this process, cross-grade items that had been selected as exemplars by two grade groups (three such items were selected by grade groups 4 and 8) were assigned to one grade by whole-group consensus. In addition, items were evaluated by the whole group for overall quality. Two items were rejected by the group during this process due to possible bias. This process yielded 14 items as recommended exemplars for grade 4, 11 items as recommended exemplars for grade 8, and 14 items as recommended exemplars for grade 12.

Mapping Panelists' Ratings to the NAEP Scales

The process of mapping panelists' ratings to the NAEP scales made significant use of *item response theory* (IRT). IRT provides statistically sophisticated methods for determining the expected performance of examinees on particular test items in terms of an appropriate measurement scale. The same measurement scale simultaneously describes the characteristics of the test items and the performance of the examinees. Once the item characteristics are set, it is possible to precisely determine how examinees are likely to perform on the test items at different points of the measurement scale.

The panelists' ratings of the NAEP test items were likewise linked, by definition, to the expected performance of examinees at the theoretical achievement level cut points. It was therefore feasible to use the IRT item characteristics to calculate the values on the measurement scale corresponding to each achievement level. This was done by averaging the item ratings over panelists for each achievement level and then simply using the item characteristics to find the corresponding achievement level cut points on the IRT measurement scale. This process was repeated for each of the NAEP content areas within each grade (4, 8, and 12).

In the final stage in the mapping process, the achievement level cut points on the IRT measurement scale were combined over content areas and rescaled to the NAEP score scale. Weighted averages of the achievement level cut points were computed. The weighting constants accounted for the measurement precision of the test items evaluated by the panelists, the proportion of items belonging to each NAEP content area, and the linear NAEP scale transformation. These weighted averages produced the final cut points for the Basic, Proficient, and Advanced achievement levels within each grade.

Figure G-1
Final Description of 1992 Mathematics Achievement Levels

GRADE 4

The NAEP content areas include: (1) numbers and operations; (2) measurement; (3) geometry; (4) data analysis, statistics, and probability; (5) algebra and functions. (Note: At the fourth-grade level, algebra and functions are treated in informal and exploratory ways, often through the study of patterns.) Skills are cumulative across levels—from Basic to Proficient to Advanced.

BASIC. Fourth-grade students performing at the **basic level** *should show some evidence of understanding the mathematical concepts and procedures in the five NAEP content areas.*

Specifically, fourth graders performing at the basic level should be able to estimate and use basic facts to perform simple computations with whole numbers; show some understanding of fractions and decimals; and solve simple real-world problems in all NAEP content areas. Students at this level should be able to use—though not always accurately—four-function calculators, rulers, and geometric shapes. Their written responses are often minimal and presented without supporting information.

PROFICIENT. Fourth-grade students performing at the **proficient level** *should consistently apply integrated procedural knowledge and conceptual understanding to problem solving in the five NAEP content areas.*

Specifically, fourth graders performing at the proficient level should be able to use whole numbers to estimate, compute, and determine whether results are reasonable. They should have a conceptual understanding of fractions and decimals; be able to solve real-world problems in all NAEP content areas; and use four-function calculators, rulers, and geometric shapes appropriately. Students performing at the proficient level should employ problem-solving strategies such as identifying and using appropriate information. Their written solutions should be organized and presented both with supporting information and explanations of how they were achieved.

ADVANCED. Fourth-grade students performing at the **advanced level** *should apply integrated procedural knowledge and conceptual understanding to complex and nonroutine real-world problem solving in the five NAEP content areas.*

Specifically, fourth graders performing at the advanced level should be able to solve complex and nonroutine real-world problems in all NAEP content areas. They should display mastery in the use of four-function calculators, rulers, and geometric shapes. These students are expected to draw logical conclusions and justify answers and solution processes by explaining why, as well as how, they were achieved. They should go beyond the obvious in their interpretations and be able to communicate their thoughts clearly and concisely.

Figure G-1 (continued)
Final Description of 1992 Mathematics Achievement Levels

GRADE 8

NAEP content areas: (1) numbers and operations; (2) measurement; (3) geometry; (4) data analysis, statistics, and probability; (5) algebra and functions. Skills are cumulative across all levels—from Basic to Proficient to Advanced.

BASIC. Eighth-grade students performing at the **basic level** *should exhibit evidence of conceptual and procedural understanding in the five NAEP content areas*. This level of performance signifies an understanding of arithmetic operations—including estimation—on whole numbers, decimals, fractions, and percents.

Eighth graders performing at the basic level should complete problems correctly with the help of structural prompts such as diagrams, charts, and graphs. They should be able to solve problems in all NAEP content areas through the appropriate selection and use of strategies and technological tools—including calculators, computers, and geometric shapes. Students at this level also should be able to use fundamental algebraic and informal geometric concepts in problem solving.

As they approach the proficient level, students at the basic level should be able to determine which of available data are necessary and sufficient for correct solutions and use them in problem solving. However, these eighth graders show limited skill in communicating mathematically.

PROFICIENT. Eighth-grade students performing at the **proficient level** *should apply mathematical concepts and procedures consistently to complex problems in the five NAEP content areas*.

They should be able to conjecture, defend their ideas, and give supporting examples. They should understand the connections between fractions, percents, decimals, and other mathematical topics such as algebra and functions. Students at this level are expected to have a thorough understanding of basic-level arithmetic operations—an understanding sufficient for problem solving in practical situations.

Quantity and spatial relationships in problem solving and reasoning should be familiar to them, and they should be able to convey underlying reasoning skills beyond the level of arithmetic. They should be able to compare and contrast mathematical ideas and generate their own examples. These students should make inferences from data and graphs; apply properties of informal geometry; and accurately use the tools of technology. Students at this level should understand the process of gathering and organizing data and be able to calculate, evaluate, and communicate results within the domain of statistics and probability.

ADVANCED. Eighth-grade students at the **advanced level** *should be able to reach beyond the recognition, identification, and application of mathematical rules in order to generalize and synthesize concepts and principles in the five NAEP content areas*.

They should be able to probe examples and counter examples in order to shape generalizations from which they can develop models. Eighth graders performing at the advanced level should use number sense and geometric awareness to consider the reasonableness of an answer. They are expected to use abstract thinking to create unique problem-solving techniques and explain the reasoning processes underlying their conclusions.

Figure G-1 (continued)
Final Description of 1992 Mathematics Achievement Levels

GRADE 12

NAEP content areas: (1) numbers and operations; (2) measurement; (3) geometry; (4) data analysis, statistics, and probability; (5) algebra and functions. Skills are cumulative across levels—from Basic to Proficient to Advanced.

BASIC. Twelfth-grade students at the **basic level** *should demonstrate procedural and conceptual knowledge in solving problems in the five NAEP content areas.*

They should be able to use estimation to verify solutions and determine the reasonableness of results as applied to real-world problems. They are expected to use algebraic and geometric reasoning strategies to solve problems. Twelfth graders performing at the basic level should recognize relationships presented in verbal, algebraic, tabular, and graphical forms; and demonstrate knowledge of geometric relationships and corresponding measurement skills.

Twelfth graders at the basic level should be able to apply statistical reasoning in the organization and display of data and in reading tables and graphs. They also should be able to generalize from patterns and examples in the areas of algebra, geometry, and statistics. At this level, they should use correct mathematical language and symbols to communicate mathematical relationships and reasoning processes; and use calculators appropriately to solve problems.

PROFICIENT. Twelfth-grade students at the **proficient level** *should consistently integrate mathematical concepts and procedures to the solutions of more complex problems in the five NAEP content areas.*

Twelfth graders performing at the proficient level should demonstrate an understanding of algebraic, statistical, and geometric and spatial reasoning. They should be able to perform algebraic operations involving polynomials; justify geometric relationships; and judge and defend the reasonableness of answers as applied to real-world situations. These students should be able to analyze and interpret data in tabular and graphical form; understand and use elements of the function concept in symbolic, graphical, and tabular form; and make conjectures, defend ideas, and give supporting examples.

ADVANCED. Twelfth-grade students at the **advanced level** *should consistently demonstrate the integration of procedural and conceptual knowledge and the synthesis of ideas in the five NAEP content areas.*

They should understand the function concept; and be able to compare and apply the numeric, algebraic, and graphical properties of functions. They should apply their knowledge of algebra, geometry, and statistics to solve problems in more advanced areas of continuous and discrete mathematics.

Twelfth graders performing at the advanced level should be able to formulate generalizations and create models through probing examples and counterexamples. They are expected to communicate their mathematical reasoning through the clear, concise, and correct use of mathematical symbolism and logical thinking.

Figure G-2
Draft Descriptions of the Achievement Levels
Prepared by the Original Level-setting Panel

Fourth-grade Draft Descriptions

BASIC. The Basic level signifies some evidence of conceptual and procedural understanding in the five NAEP content areas of Numbers and Operations; Measurement; Geometry; Data Analysis, Statistics, and Probability; and Algebra and Functions. Understanding simple facts and single-step operations are included at this level, as is the ability to perform simple computations with whole numbers. This level shows a partial mastery of estimation, basic fractions, and decimals relating to money or the number line; it shows an ability to solve simple real-world problems involving measurement, probability, statistics, and geometry. At this level, there is a partial mastery of tools such as four-function calculators and manipulatives (geometric shapes and rulers). Written responses are often minimal, perhaps with a partial response and lack of supportive information.

PROFICIENT. The Proficient level signifies consistent demonstration of the integration of procedural knowledge and conceptual understanding as applied to problem solving in the five NAEP content areas of Numbers and Operations; Measurement; Geometry; Data Analysis, Statistics, and Probability; and Algebra and Functions. The Proficient level indicates an ability to perform computation and estimation with whole numbers, to identify fractions, and to work with decimals involving money or the number line. Solving real-world problems involving measurement, probability, statistics, and geometry is an important part of this level. This level signifies the ability to use, as tools, four-function calculators, rulers, and manipulatives (geometric shapes). It includes the ability to identify and use pertinent/appropriate information in problem settings. The ability to make connections between and among skills and concepts emerges at this level. Clear and organized written presentations, with supportive information, are typical. And, there is an ability to explain how the solution was achieved.

ADVANCED. The Advanced level signifies the integration of procedural knowledge and conceptual understanding as applied to problem solving in the five NAEP content areas of Numbers and Operations; Measurement; Geometry; Data Analysis, Statistics, and Probability; and Algebra and Functions. This is evidenced by divergent and elaborate written responses. The Advanced level indicates an ability to solve multistep and nonroutine real-world problems involving measurement, probability, statistics, and geometry, and an ability to perform complex tasks involving multiple steps and variables. Tools are mastered, including four-function calculators, rulers, and manipulatives (geometric shapes). This level signifies the ability to apply facts and procedures by explaining *why* as well as *how*. Interpretations extend beyond obvious connections and thoughts are communicated clearly and concisely. At this level, logical conclusions can be drawn and complete justifications can be provided for answers and/or solution processes.

Eighth-grade Draft Descriptions

BASIC. Basic students should begin to describe objects, to process accurately and elaborate relationships, to compare and contrast, to find patterns, to reason from graphs, and to understand spatial reasoning. This level of partial mastery signifies an understanding of arithmetic operations on whole numbers, decimals, fractions, and percents, including estimation. Problems that are already set up are generally solved correctly, as are one-step problems. However, problems involving the use of available data, and determinations of what is necessary and sufficient to solve the problem, are generally quite difficult. Students should select appropriate problem-solving tools, including calculators, computers, and manipulatives (geometric shapes) to solve problems from the five content areas. Students should also be able to use elementary algebraic concepts

Figure G-2 (continued)
Draft Descriptions of the Achievement Levels
Prepared by the Original Level-setting Panel

and elementary geometric concepts to solve problems. This level indicates familiarity with the general characteristics of measurement. Students at this level may demonstrate limited ability to communicate mathematical ideas.

PROFICIENT. Proficient students apply mathematical concepts consistently to more complex problems. They should make conjectures, defend their ideas, and give supporting examples. They have developed the ability to relate the connections between fractions, percents, and decimals, as well as other mathematical topics. The Proficient level denotes a thorough understanding of the arithmetic operations listed at the Basic level. This understanding is sufficient to permit applications to problem solving in practical situations. Quantity and spatial relationships are familiar situations for problem solving and reasoning, and this level signifies an ability to convey the underlying reasoning skills beyond the level of arithmetic. Ability to compare and contrast mathematical ideas and generating examples is within the Proficient domain. Proficient students can make inferences from data and graphs; they understand the process of gathering and organizing data, calculating and evaluating within the domain of statistics and probability, and communicating the results. The Proficient level includes the ability to apply the properties of elementary geometry. Students at this level should accurately use the appropriate tools of technology.

ADVANCED. The Advanced level is characterized by the ability to go beyond recognition, identification, and application of mathematical rules in order to generalize and synthesize concepts and principles. Generalization often takes shape through probing examples and counterexamples and can be focused toward creating models. Mathematical concepts and relationships are frequently communicated with mathematical language, using symbolic representations where appropriate. Students at the Advanced level consider the reasonableness of an answer, with both number sense and geometric awareness. Their abstract thinking ability allows them to create unique problem-solving techniques and explain the reasoning processes they followed in reaching a conclusion. These students can probe through examples and counterexamples that allow generalization and description of assumptions with models and elegant mathematical language.

Twelfth-grade Draft Descriptions

BASIC. This level represents understanding of fundamental algebraic operations with real numbers, including the ability to solve two-step computational problems. It also signifies an understanding of elementary geometrical concepts such as area, perimeter, and volume, and the ability to make measurements of length, weight, capacity, and time. Also included at the Basic level is the ability to comprehend data in both tabular and graphical form and to translate between verbal, algebraic, and graphical forms of linear expression. Students at this level should be able to use a calculator appropriately.

PROFICIENT. This level represents mastery of fundamental algebraic operations and concepts with real numbers, and an understanding of complex numbers. It also represents understanding of polynomials and their graphs up to the second degree, including conic sections. The elements of plane, solid, and coordinate geometry should be understood at the Proficient level. The Proficient level includes the ability to apply concepts and formulas to problem solving. Students at this level should demonstrate critical thinking skills. The Proficient level also represents the ability to judge the reasonableness of answers and the ability to analyze and interpret

Figure G-2 (continued)
Draft Descriptions of the Achievement Levels
Prepared by the Original Level-setting Panel

data in both tabular and graphical form. Basic algebraic concepts, measurement, and constructive geometry concepts are mastered at this level.

ADVANCED. The Advanced level represents mastery of trigonometric, exponential, logarithmic, and composite functions, zeros and inverses of functions, polynomials of the third degree and higher, rational functions, and graphs of all of these. In addition, the Advanced level represents mastery of topics in discrete mathematics including matrices and determinants, sequences and series, and probability and statistics, as well as topics in analytic geometry. The Advanced level also signifies the ability to successfully apply these concepts to a variety of problem-solving situations.

Figure G-3
Revised Draft Descriptions of the Achievement Levels
Recommended by the Follow-Up Validation Panel

Revised Fourth-grade Draft Descriptions

BASIC. Basic students exhibit some evidence of conceptual and procedure understanding in the five NAEP content areas. At the fourth grade level, algebra and functions are treated in informal and exploratory ways often through the study of patterns. Basic students estimate and use basic facts to perform simple computations with whole numbers. These students show some understanding of fractions and decimals. They solve simple real world problems in all areas. These students use, although not always accurately, four-function calculators, rulers, and geometric shapes. Written responses are often minimal and lack supporting information.

PROFICIENT. Proficient students consistently integrate procedural knowledge and conceptual understanding as applied to problem solving in the five NAEP content areas. Using whole numbers they estimate, compute, and determine whether their results are reasonable. They have a conceptual understanding of fractions and decimals. Solving real world problems in all areas is important at this level. Proficient students appropriately use four-function calculators, rulers and geometric shapes. These students use problem solving strategies such as identifying and using appropriate information. [Problem-solving strategies include identification and use of appropriate information.] They present organized written solutions with supporting information and explain how they were achieved.

ADVANCED. Advanced students integrate procedural knowledge and conceptual understanding as applied to problem solving in the five NAEP content areas. They solve complex and non-routine real world problems in all areas. They have mastered the use of tools such as four-function calculators, rulers and geometric shapes. Advanced students draw logical conclusions and justify answers and solution processes by explaining the "why" as well as the "how." Interpretations extend beyond obvious connections and thoughts are communicated clearly and concisely.

Revised Eighth-grade Draft Descriptions

BASIC. Basic students exhibit evidence of conceptual and procedural understanding. These students compare and contrast, find patterns, reason from graphs, and understand spatial reasoning. This level of performance signifies an understanding of arithmetic operations, including estimation, on whole numbers, decimals, fractions, and percents. Students complete problems correctly with the help of structural prompts such as diagrams, charts, and graphs. As students approach the proficient level, they will solve problems involving the use of available data, and determine what is necessary and sufficient for a correct solution. Students use problem solving strategies and select appropriate tools, including calculators, computers, and manipulatives (geometric shapes) to solve problems from the five content areas. Students use fundamental algebraic and informal geometric concepts to solve problems. Students at this level demonstrate limited skills in communicating mathematically.

PROFICIENT. Proficient students apply mathematical concepts and procedures consistently to complex problems. They make conjectures, defend their ideas, and give supporting examples. They have developed the ability to relate the connections between fractions, percents, and decimals, as well as other mathematical topics, such as algebra and functions. The proficient level denotes a thorough understanding of the arithmetic

Figure G-3 (continued)
Revised Draft Descriptions of the Achievement Levels
Recommended by the Follow-Up Validation Panel

operations listed at the basic level. This understanding is sufficient to permit applications to problem solving in practical situations. Quantity and spatial relationships are familiar situations for problem solving and reasoning, and students at this level convey the underlying reasoning skills beyond the level of arithmetic. Proficient students compare and contrast mathematical ideas and generate their own examples. These students make inferences from data and graphs; they understand the process of gathering and organizing data, calculating, evaluating, and communicating the results within the domain of statistics and probability. Proficient students apply the properties of informal geometry, and accurately use the appropriate tools of technology.

ADVANCED. Advanced students go beyond recognition, identification, and application of mathematical rules in order to generalize and synthesize concepts and principles. Generalization often takes shape through probing examples and counter examples and can be used to create models. Mathematical concepts and relationships are frequently communicated with mathematical language, using symbolic representations where appropriate. Students at the advanced level consider the reasonableness of an answer, with both number sense and geometric awareness. Their abstract thinking allows them to create unique problem solving techniques and explain the reasoning processes they followed in reaching a conclusion. These students probe examples and counter examples that allow generalization and description of assumptions with models and elegant mathematical language.

Revised Twelfth-grade Draft Descriptions

BASIC. Basic students demonstrate procedural and conceptual knowledge in solving problems in the five NAEP content areas. They use estimation to verify solutions and determine the reasonableness of the results to real world problems. Algebraic and geometric reasoning strategies are used to solve problems. These students recognize relationships in verbal, algebraic, tabular, and graphical forms. Basic students demonstrate knowledge of geometric relationships as well as corresponding measurement skills. Statistical reasoning is applied to the organization and display of data and to reading tables and graphs. These students generalize from patterns and examples in the areas of algebra, geometry, and statistics. They communicate mathematical relationships and reasoning processes with correct mathematical language and symbolic representations. Calculators are used appropriately to solve problems.

PROFICIENT. Proficient students integrate mathematical concepts and procedures consistently to more complex problems in the five NAEP content areas. They demonstrate an understanding of algebraic reasoning, geometric and spatial reasoning, and statistical reasoning as applied to other areas of mathematics. They perform algebraic operations involving polynomials, justify geometric relationships, and judge and defend the reasonableness of answers in real world situations. These students analyze and interpret data in tabular and graphical form. Proficient students understand and use elements of the function concept in symbolic, graphical and tabular form. They make conjectures, defend their ideas, and give supporting examples.

ADVANCED. Advanced students consistently demonstrate the integration of procedural and conceptual knowledge, as well as the synthesis of ideas, in the five NAEP content areas. Advanced students understand the function concept, and they compare and apply the numeric. algebraic, and graphical properties of functions. They apply and connect their knowledge of algebra, geometry, and statistics to solve problems in more advanced areas of continuous and discrete mathematics. Advanced students formulate generalizations using examples and counter examples to create models. In communicating their mathematical reasoning, these students demonstrate clear, concise, and correct use of mathematical symbolism and logical thinking.

Figure G-4
Meeting Participants, NAEP Mathematics Achievement Level-setting
Original Meeting, St. Louis, Missouri, March 20-24, 1992

Marge Blizard
Blizard Professional Cleaning
Franklin, Connecticut

Christopher Chomyak
The Episcopal Church
Calais, Maine

Janet Green
Met Life
Crownsville, Maryland

Mary Norman
Dekalb County Board of Education
Decatur, Georgia

Janice Wamsley
Alcorn School System
Glen, Mississippi

Ronald Higgins
Walla Walla School District
Walla Walla, Washington

Leona Lee
Baltimore City Public Schools
Baltimore, Maryland

Lisa Bietau
USD 383 Manhattan Public Schools
Manhattan, Kansas

Marsha Davis
Alcorn County Schools
Corinth, Mississippi

Jean Bush Ragin
Patterson High School
Baltimore, Maryland

Bill Oldham
Harding University
Searcy, Arkansas

George Shell
Retired Principal
Draper, Utah

Marsha Stovey
Detroit Public Schools
Detroit, Michigan

Vance Morris
Dekalb County Board of Education
Atlanta, Georgia

Carol Ballentine
Duval County Schools
Jacksonville, Florida

Tami Harvey, ESD
Audiometric Technician
Burns, Oregon

Laurence Payne
Greater Houston Coalition
for Educational Excellence
Houston, Texas

Cheryl Yunk
USD 383
Manhattan, Kansas

Kirby Gchachu
Zuni Public School District
Zuni, New Mexico

Corliss Hubert
Rutherford Board of Education
Englewood, New Jersey

Joyce Dunn
Alcorn County Schools
Corinth, Mississippi

Gloria Moran
Williams Junior High
Bridgewater, Massachusetts

Charles Jackson
Blairsville, Pennsylvania

Cassandra Turner
Internal Revenue Service
Miami, Florida

Figure G-4 (continued)
Meeting Participants, NAEP Mathematics Achievement Level-setting
Original Meeting, St. Louis, Missouri, March 20-24, 1992

Jack Deal
Bethel Park School District
Pittsburgh, Pennsylvania

Ninfa Rivera
Lyford CISD
Raymondville, Texas

Gerald Zeringue
Garrity Construction Company
Harvey, Louisiana

Linda Brown
Van Zile Elementary School
Detroit, Michigan

Judy Bibb
Lonoke High School
Cabot, Arkansas

David Rank
School District of Greenville
Greenville, South Carolina

John Sweeney
Freed-Haideman University
Henderson, Tennessee

Nancy Pejouhy
Woodstock Union High School
Woodstock, Vermont

Jim Trefzger
Parkland College
Champaign, Illinois

Joanne Greaver
Jefferson County Public Schools
Louisville, Kentucky

Ellie Cucinatto
Bridgewater Public Schools
Bridgewater, Massachusetts

Lillie Carr
Pender County Schools
Teachey, North Carolina

Eric Cain
IBM
Metairie, Louisiana

Phillip Stroup
Butler County MR/DD
Seven Mile, Ohio

Mike Gobel
Walla Walla School District
Walla Walla, Washington

Juanita Tietze
Retired Principal
Canton, Ohio

Bill Cramer, Jr.
Cramer & Mallon, Attorneys at Law
Burns, Oregon

Norma Newman
Ysleta Independent School District
El Paso, Texas

William Rickenbach
Bethel Park School District
Bethel Park, Pennsylvania

Violet Cosgrove
Retired
Glen Burnie, Maryland

Danny McDougal
Pre-Mc, Inc.
Allen, Oklahoma

Bill Anderson
Administration Eagle Union
Zionsville, Indiana

Dan Thompson
Thompson Construction Company
Trinidad, Colorado

Figure G-4 (continued)
Meeting Participants, NAEP Mathematics Achievement Level-setting
Original Meeting, St. Louis, Missouri, March 20-24, 1992

Nancy Gallagher
West Penn Power Company
Kittanning, Pennsylvania

William Hawes
The Hawes Company
Tucker, Georgia

Zhining Qin
Minnesota Department of Education
St. Paul, Minnesota

Charles McGee
Greenville County School District
Greenville, South Carolina

Barbara Bayne
Greenville County School District
Greenville, South Carolina

Landa McLaurin
Baltimore City Schools
Baltimore, Maryland

Nancy Potempa
St. Xavier University
Mokena, Illinois

Florencetine Jasmin
Baltimore City Public Schools
Baltimore, Maryland

Jeane Joyner
Dept. of Public Instruction
Raleigh, North Carolina

Carolyn Craig
State Department of Education
Jackson, Mississippi

Mary Bennion
SFUSD
San Francisco, California

Bill Eyestone
Walla Walla School District
Walla Walla, Washington

Mary Gahn
New York Board of Education
Westbury, New York

Barbara Faltz-Jackson
Baltimore City Public Schools
Baltimore, Maryland

Florence Kelly
Manville Board of Education
Manville, New Jersey

Philip Brach
Univ. of the District of Columbia
Washington, D.C.

Larry Brown
Oil industry (Self-Employed)
Allen, Oklahoma

W. Garry Quast
Slippery Rock University
Slippery Rock, Pennsylvania

Carl Springfels
Consultant (Self-Employed)
Miami Shores, Florida

Anna Maria Golan
Santa Ana Unified
Fountain Valley, California

Ricardo Suarez
Lyford CISD
Raymondville, Texas

Figure G-5
Meeting Participants, NAEP Mathematics Achievement Level-setting
Follow-up Validation Meeting, Nantucket, Massachusetts, July 17-19, 1992

Charles Allen
Michigan Department of Education
Lansing, Michigan

Linda Brown
Van Zile Elementary School
Clinton Township, Michigan

Ellie Cucinatto
Bridgewater Public Schools
Bridgewater, Massachusetts

Jack Deal
Bethel Park School District
Pittsburgh, Pennsylvania

Paula Duckett
River Terrace Community School Board
Washington, DC

Edward Esty
SRI International
Washington, D.C.

Barbara Faltz-Jackson
Baltimore Public Schools
Baltimore, Maryland

Joan Ferini-Mundy
University of New Hampshire
Durham, New Hampshire

Marilyn Hala
National Council of Teachers of Mathematics
Washington, D.C.

Florence Kelly
Largo Public Schools
Largo, Florida

Henry Kepner
University of Wisconsin at Milwaukee
Milwaukee, Wisconsin

Charles McGee
Greenville Public Schools
Greenville, South Carolina

Landa McLaurin
Baltimore City Schools
Baltimore, Maryland

Gloria Moran
Williams Junior High School
Bridgewater, Massachusetts

Jo Ann Mosier
Kentucky Department of Education
Frankfort, Kentucky

Mary Norman
DeKalb County Board of Education
Decatur, Georgia

David Rank
Greenville Public Schools
Greenville, South Carolina

Sharon Steglein
Minnesota Department of Education
St. Paul, Minnesota

APPENDIX H

THE NAEP ACHIEVEMENT LEVEL-SETTING PROCESS FOR THE 1992 READING ASSESSMENT

APPENDIX H

The NAEP Achievement Level-setting Process for the 1992 Reading Assessment[1]

Mary Lyn Bourque[2]
National Assessment Governing Board

Introduction

Since 1984, NAEP has reported the performance of students in the nation and for specific subpopulations on a 0-to-500 proficiency scale. The history and development of the scale and the anchoring procedure used to interpret specific points on that scale are described in Appendix G.

Legislation[3] in 1988 created an independent board, the National Assessment Governing Board (NAGB), responsible for setting policy for the NAEP program. The Board has a statutory mandate to identify "appropriate achievement goals for each...grade in each subject area to be tested under the National Assessment." Consistent with this directive, and striving to achieve one of the primary mandates of the statute "to improve the form and use of NAEP results," the Board set performance standards (called achievement levels by NAGB) on the National Assessment in 1990, and again in 1992.

The 1990 trial, initiated in December 1989 with the dissemination of a draft policy statement (NAGB, 1989) and culminating 22 months later in the publication of the NAGB report, *The Levels of Mathematics Achievement* (Bourque & Garrison, 1991), consisted of two phases: the main study and a replication-validation study. Although there were slight differences between the two phases, there were many common elements. Both phases used a modified (iterative/empirical) Angoff (1971) procedure for arriving at the levels; both focused on estimating performance levels based on a review of the 1990 NAEP mathematics item pool;

---

[1] Several studies have examined the validity of the achievement level-setting process, including: 1) American College Testing. (1992). *Setting achievement levels on the 1992 National Assessment of Education Progress in mathematics, reading, and writing: A technical report on the reliability and validity.* Iowa City, IA: Author. 2) Burstein, L., Koretz, D. M., Linn, R. L., Sugrue, B., Novak, J., Lewis, E., & Baker, E. L. (1993). *The validity of interpretations of the 1992 NAEP achievement levels in mathematics.* Los Angeles, CA: CRESST. 3) National Academy of Education. (1993). *Setting performance standards for student achievement.* Stanford, CA: Author. 4) National Academy of Education. (1994). *Setting performance standards for student achievement: Background studies.* Stanford, CA: Author. 5) National Academy of Education. (1993). *The Trial State Assessment: Prospects and realities.* Stanford, CA: Author.

[2] The author is grateful to Susan Loomis, Richard Luecht, and Mel Webb, American College Testing, for their helpful suggestions and comments for improving earlier drafts of this paper.

[3] Public Law 100-297. (1988). National Assessment of Educational Progress improvement act (Article No. USC 1221). Washington, DC.

and both phases employed policy definitions for basic, proficient, and advanced levels (NAGB, 1990) as the criteria for rating items. The 1990 process was evaluated by a number of different groups (for a discussion, see Hambleton & Bourque, 1991) who identified technical flaws in the 1990 process. These evaluations influenced the Board's decision to set the levels again in 1992, and to not use the 1990 levels as benchmarks for progress toward the national goals during the coming decade. It is interesting to note, however, that the 1990 and 1992 processes produced remarkably similar results.

In September 1991, the Board contracted with American College Testing (ACT) to convene the panels of judges that would recommend the levels on the 1992 NAEP assessments in reading, writing, and mathematics. While the 1992 level-setting activities were not unlike those undertaken by the Board in 1990, there were significant improvements made in the process for 1992. There was a concerted effort to bring greater technical expertise to the process: the contractor selected by the Board has a national reputation for setting standards in a large number of certification and licensure exams; an internal and external advisory team monitored all the technical decisions made by the contractor throughout the process; and state assessment directors periodically provided their expertise and technical assistance at key stages in the project.

Setting achievement levels is a method for setting standards on the NAEP assessment that identify what students should know and be able to do at various points along the proficiency scale. The initial policy definitions of the achievement levels were presented to panelists along with an illustrative framework for more in-depth development and operationalization of the levels. Panelists were asked to determine descriptions/definitions of the three levels from the specific framework developed for the NAEP assessment with respect to the content and skills to be assessed. The operationalized definitions were refined throughout the level-setting process, as well as validated with a supplementary group of judges subsequent to the level-setting meetings. Panelists were also asked to develop a list of illustrative tasks associated with each of the levels, after which sample items from the NAEP item pool were identified to exemplify the full range of performance of the intervals between levels. The emphasis in operationalizing the definitions and in identifying and selecting exemplar items and papers was to represent the full range of performance from the lower level to the next higher level. The details of the implementation procedures are outlined in the remainder of this appendix.

Preparing for the Reading Level-setting Meeting

It is important for the planning of any standard-setting effort to know how various process elements interact with each other. For example, panelists interact with pre-meeting materials, meeting materials (i.e., the assessment questions, rating forms, rater feedback, and so forth), each other, and the project staff. All of these elements combine to promote or degrade what has been called intrajudge consistency and interjudge consensus (Friedman & Ho, 1990).

Previous research has conceptualized the effects of two major kinds of interaction: (1) people interacting with text (Smith & Smith, 1988), and (2) people interacting with each other Curry, 1987; Fitzpatrick, 1989). In order to assess the effects of textual and social interaction and adjust the standard-setting procedures accordingly, a pilot study was conducted as the first phase of the 1992 initiative.

Reading was chosen as the single content area to be pilot tested since it combined all of the various features found in the other NAEP assessments, including multiple-choice, and both short and extended constructed-response items. The pilot study provided the opportunity to implement and evaluate all aspects of the operational plan—background materials, meeting materials, study design, meeting logistics, staff function, and participant function.

The overall pilot was quite successful. The level-setting process worked well, and the pilot allowed the contractor to make improvements in the design before implementation activities began. For example, schedule changes were made that allowed the panelists more time to operationalize the policy definitions before beginning the item-rating task. Also, the feedback mechanisms used to inform panelists about interjudge and intrajudge consistency data were improved for clarity and utility to the entire process.

The Reading Level-setting Panel

Sixty-four panelists representing 32 jurisdictions (31 states and the Virgin Islands) were selected from the 366 nominees and invited to participate in the level-setting process. They represented reading/language arts teachers at grades 4, 8, and 12, nonteacher educators, and members of the noneducator (general public) community. The group was balanced by gender, race/ethnicity, NAEP regions of the country, community type (low SES, not low SES), district size, and school type (public/private). Two panelists were unable to attend due to a family emergency and a loss of job, resulting in 62 participants, 22 at grade 4, 20 at grade 8, and 20 at grade 12.

Process for Developing the Achievement Levels

The four-and-one-half-day session began with a brief overview of NAEP and NAGB, a presentation on the policy definitions of the achievement levels, a review of the NAEP reading assessment framework, and a discussion of factors that influence item difficulty. The purpose of the presentation was to focus panelists' attention on the reading framework and to emphasize the fact that panelists' work was directly related to the NAEP assessment, not to the whole domain of reading.

All panelists completed and self-scored an appropriate grade-level form of the NAEP assessment. The purpose of this exercise was to familiarize panelists with the test content and scoring protocols—as well as time constraints—before beginning to develop the preliminary operationalized descriptions of the three levels.

Working in small groups of five or six, then eventually in grade-level groups, panelists expanded and operationalized the policy definitions of basic, proficient, and advanced in terms of specific reading skills, knowledge, and behaviors that were judged to be appropriate expectations for students in each grade, and to be in accordance with the current reading assessment framework.

The policy definitions are as follows:

Basic	This level, below proficient, denotes partial mastery of the knowledge and skills that are fundamental for proficient work at each grade—4, 8, and 12.
Proficient	This central level represents solid academic performance for each grade tested—4, 8, and 12. Students reaching this level have demonstrated competency over challenging subject matter and are well prepared for the next level of schooling.
Advanced	This higher level signifies superior performance beyond proficient grade-level mastery at grades 4, 8, and 12.

The small groups were allowed to brainstorm about what student performance *should be*, using the framework and their experience in completing the NAEP assessment as guides[4]. In addition, a practice task caused panelists to examine items in the half of the item pool that they would not be rating later. A comprehensive listing of grade level descriptors was developed, and panelists were asked to identify the five or six that best described what students *should be able to do* at each of the levels. Those descriptors appearing with the greatest frequency were compiled into a discussion list for the grade-level groups. Additions, deletions, and modifications were made as a result of discussions, and the groups reached general agreement that the final list of descriptors represented what students *should be able to do* at each achievement level.

Panelists next received two hours of training in the Angoff method. Training was customized to reflect the unique item formats of the particular subject area assessment. Once a conceptual consensus was reached about the characteristics of **marginally** acceptable examinees at each of the three levels, practice items from the released pool were rated by the panelists according to the process defined in the contractor's plan. For multiple-choice and short constructed-response items, panelists were asked to rate each item for the expected probability of a correct response for a group of *marginally* acceptable examinees at the basic, proficient, and advanced levels. For extended constructed-response items, panelists were asked to review a set of student response papers and select three papers, one for each achievement level, that typified *marginally* acceptable examinee performance for that level.

Following training in the Angoff method, the judges began the rating and paper selection process, inspecting and rating each dichotomously scored item in the pool for the expected probabilities of answering the item correctly at each level. For polytomously scored items, panelists reviewed a representative set of 24 to 28 student response papers for each item and selected the paper that best represented marginally acceptable student performance at each level. Panelists completed three rounds of item ratings and paper selections. For Round 1, panelists first answered the items related to a reading passage, then reviewed their answers using scoring keys and protocols. This process helped ensure that panelists would be thoroughly familiar with each item, including the foils and scoring rubrics, before rating the item. Panelists provided item ratings/paper selections for all three achievement levels, one item at a time, for all the items related to a reading passage, then proceeded to the next reading passage and set of items, for which the process was repeated. Panelists rated items for half the items in their grade-level assessment; one block of exercises was common to both halves of the grade-level groups. During Round 1, panelists used their lists of descriptors and other training materials for guidance in the rating process.

---

[4] The panelists also reviewed about half the item pool (the half they would not be rating later) so that the descriptors could be further modified if that was deemed appropriate.

Following Round 1, item response theory (IRT) was used to convert the rating results[5] for each rater to a latent ability scale, represented by the Greek letter theta (θ). This θ scale was the same scale to which the NAEP items evaluated by each panelist were calibrated. In order to provide meaningful feedback about item ratings, a special *relative scale* was constructed, which was a linear transformation of the theta scale having a mean of 75 and standard deviation of 15. Before Round 2 of the rating process, panelists were given interjudge consistency information using this relative scale. This information allowed panelists to see where their individual mean item ratings were on the scale, relative to the mean for the group and to the means for other panelists. Reasons for extreme mean ratings, including the possibility that some panelists misinterpreted the item rating task, were discussed briefly.

Before Round 2, panelists were also given item difficulty data. This information was presented as the percentage of students who answered each item correctly during the actual NAEP administration, for items scored "correct" or "incorrect" (i.e., multiple-choice and short constructed-response items), and as the mean score for student responses (on a scale of 1 to 4) for the extended response items. Panelists were told that this item difficulty information should be used as a reality check. For items on which item ratings differed substantially from the item difficulty value, panelists were asked to reexamine the item to determine if they had misinterpreted the item or misjudged its difficulty. Results of the data analysis, and panelists' own evaluations, indicated that the item difficulty information was perceived as very useful but had little impact on panelists' ratings.

For Round 2, panelists reviewed the same set of items they rated in Round 1 and, using the interjudge consistency information, the item difficulty information, and the information provided prior to Round 1, they either confirmed their initial item ratings and paper selections or adjusted their ratings to reflect the additional information. About one-half of Round 1 item ratings and paper selections were adjusted during Round 2.

Prior to Round 3, panelists' ratings were reanalyzed and additional information was presented to panelists concerning intrajudge variability. For each panelist, the intrajudge variability information consisted of those items that they had rated differently than items having similar difficulty, taking into consideration the panelist's aggregated item ratings. That is, the panelists' aggregated item ratings were converted to the theta (θ) scale. All items rated by the panelists were then analyzed in terms of the panelist's achievement level (θ) in comparison to actual student performance on the items. The observed item rating from each panelist was contrasted to an expected item rating. Those items with the largest differences between observed and expected ratings were identified. Panelists were given this information and asked to review each of these items and decide if their Round 2 ratings still accurately reflected their best judgments of the items. The intrajudge consistency data was to be used to flag items for reconsideration in the final round of rating.

---

[5] Because the IRT item parameters were not available for the polytomously scored (extended constructed-response) items, these items were not included in the following discussion of results.

For Round 3, panelists reviewed the same set of items they rated in Rounds 1 and 2 using both the new intrajudge variability information and the information made available during Rounds 1 and 2. In addition, panelists could discuss, within their small groups, ratings and paper selections for specific items about which they were unsure. About one-third of the item ratings were adjusted during Round 3.

Process of Selecting Exemplar Items

On the final day of the achievement level-setting process, panelists reviewed items from the 1992 item pool scheduled for release to the public. The released item pool was the set from which the panelists could select items illustrative of the achievement levels for their grade. Exercises are organized in blocks, consisting of a reading passage, followed by several items, usually employing each of the three item formats, (i.e., multiple-choice, short constructed-response, and extended constructed-response). A total of 10 blocks from the 1992 exercise pool were scheduled for release: 2 blocks from the fourth-grade pool, totaling 19 items; 4 blocks from the eighth-grade pool, totaling 52 items; and 4 blocks from the twelfth-grade pool, totaling 46 items.

Panelists who had rated specific blocks of released items were asked to review those same items again to select particular ones as exemplary of each achievement level. The items were pre-assigned to each achievement level based on the final round of the judges' rating data, and using the following statistical criteria. For any given level (basic, proficient, or advanced),

(1) items having an expected p-value[6] $\geq .501$ and $\leq .750$, *at that level,* were assigned to that level;

(2) items meeting the criteria at *more than one level* were assigned to *one* level taking both the expected p-value and the appropriateness of the item for one of the levels into account; and

(3) items with expected p-values $<.501$ were assigned to levels where a specific passage had few or no items at that level.

For example, the raters' expected p-value for one of the released items might have been .366 at the basic level, .701 at the proficient level, and .932 at the advanced level. This item would have been identified for review as a potential exemplar item for the proficient level. The expected p-value at the basic level was too low for consideration as a basic-level exemplar—that is, the item was judged to be too difficult, and the expected p-value at the advanced level was too high for consideration at the advanced level—that is, the item was judged to be too easy. Table H-1 shows the results of this process for each grade and level.

---

[6] Expected p-values were based on the average predicted performance at the cut point for each achievement level.

Table H-1

Results of First Review for Achievement-level Exemplars

Level/Status	Grade 4	Grade 8	Grade 12	All Grades
Total released	19	52	46	117
Basic				
Reviewed	4	12	18	34
Recommended	3	5	14	22
Proficient				
Reviewed	5	14	20	39
Recommended	4	12	9	25
Advanced				
Reviewed	5	6	7	18
Recommended	5	6	8	19

Table H-2

Results of Review of Additional Items for Achievement-level Exemplars

Level/Status	Grade 4	Grade 8	Grade 12	All Grades
Total items recommended	13	13	21	47
Basic				
Reviewed	3	12	12	27
Recommended	6	7	8	21
Proficient				
Reviewed	4	13	11	28
Recommended	6	3	8	17
Advanced				
Reviewed	5	8	9	22
Recommended	1	3	5	9

Panelists were asked to review the items as classified, and form an individual judgment regarding the suitability of each item to illustrate and further communicate the meaning of the levels. Each item's classification could be accepted, rejected, or reassigned, although the procedure was primarily designed to eliminate items that did not meet panelists' expectations for any reason. Items were reclassified if a strong consensus was found to hold for that change.

During the validation process, described in the next section, items were again reviewed. Those that had been selected by the original standard-setting panel were grouped into sets of *pre-selected* items. All remaining items in the released blocks that met the statistical criteria, *but were not recommended by the original panel*, were grouped into a set identified as *additional items for review*. Exercises that had been recommended for reclassification into another achievement level category were presented in their original classification for purposes of this review. As the Table H-2 shows, 21 items were recommended as exemplars for the basic level, 17 for the proficient level, and 9 for the advanced.

Process for Validating the Levels

Nineteen reading educators participated in the item selection and content validation process. Ten of the panelists were reading teachers who had participated in the original achievement level-setting process and who had been identified as outstanding panelists by grade group facilitators during this meeting, who were extensively involved with professional organizations (e.g., the International Reading Association, the National Reading Conference, or the National Council for Teachers of English), and who had outstanding service credentials. The other nine panelists represented state-level reading curriculum supervisors or assessment directors, as well as university faculty teaching in disciplines related to this subject area. To the extent possible, the group was balanced by race/ethnicity and gender.

The two-and-one-half day meeting began by briefing panelists on the purpose of the meeting and by giving them an overview of the level-setting process and results. Panelists first reviewed the operationalized descriptions of the achievement levels for qualities such as (1) within- and across-grade consistency, (2) grade-level appropriateness, and (3) utility for increasing the public's understanding of the NAEP reading results. Next, panelists reviewed the operationalized descriptions of the achievement levels for consistency with the NAGB policy definitions of basic, proficient, and advanced with the NAEP *Reading Objectives*. Working in grade-level (4, 8, and 12) groups of 6 to 7 panelists each, then as a whole group, panelists reviewed the operationalized descriptions to provide within- and across-grade consistency, and to align the language and concepts of the descriptions more closely with the language of the NAEP *Reading Objectives*. (Both the original descriptions and the revised descriptions are included later in this appendix.) Finally, panelists suggested revisions they thought would improve the operational descriptions based on their earlier reviews.

On the final day, panelists worked in grade-level groups to review the possible exemplar items. The task was to select a set of items, for each achievement level for their grade, that would best communicate to the public the levels of reading ability and the types of skills needed to perform in reading at that level.

After selecting sets of items for their grades, the three grade-level groups met as a whole group to review item selection. During this process, cross-grade items that had been selected as exemplars for two grades (two such items were selected for grades 8 and 12) were assigned to one grade by whole-group consensus. In addition, items were evaluated by the whole group for overall quality. This process yielded 13 items as recommended exemplars for grade 4, 13 items as recommended exemplars for grade 8, and 21 items as recommended exemplars for grade 12.

Mapping the Levels onto the NAEP Scale

The process of mapping panelists' ratings to the NAEP scales used *item response theory* (IRT). IRT provided statistically sophisticated methods for determining the expected performance of examinees on particular test items in terms of an appropriate measurement scale. The same measurement scale simultaneously described the characteristics of the test items and the performance of the examinees. Once the item characteristics were set, it was possible to determine precisely how examinees were likely to perform on the test items at different points of the measurement scale.

The panelists' ratings of the NAEP test items were likewise linked, by definition, to the expected performance of examinees at the theoretical achievement level cut points. It was therefore feasible to use the IRT item characteristics to calculate the values on the measurement scale corresponding to each achievement level. This was done by averaging the item ratings over panelists for each achievement level and then simply using the item characteristics to find the corresponding achievement level cut points on the IRT measurement scale. This process was repeated for each of the NAEP reading scales within each grade (4, 8, and 12).

For the multiple-choice and short constructed-response items that were dichotomously scored, the judges each rated half of the items in the NAEP pool in terms of the expected probability that a student at a borderline achievement level would answer the item correctly, based on the judges' operationalization of the policy definitions and the factors that influence item difficulty. To assist the judges in generating consistently scaled ratings, the rating process was repeated twice, with feedback. Information on consistency among different judges and on the difficulty of each item[7] was fed back into the first repetition (Round 2), while information on consistency within each judge's set of ratings was fed back into the second repetition (Round 3). The third round of ratings permitted the judges to discuss their ratings among themselves to resolve problematic ratings. The mean final rating of the judges aggregated across multiple-choice and short constructed-response items yielded the threshold values for these items in the percent correct metric. These cut scores were then mapped onto the NAEP scale (which is defined and scored using item response theory, rather than percent correct).

For extended constructed-response items, judges were asked to select student papers that exemplified performance at the cutpoint of each achievement level. Then for each achievement level, the mean of the scores assigned to the selected papers was mapped onto the NAEP scale in a manner similar to that used for the items scored dichotomously.

---

[7] Item difficulty estimates were based on a preliminary, partial set of responses to the national assessment.

The final cut score for each achievement level was a weighted average of the cut score for the multiple-choice and short constructed-response items and the cut score for the extended constructed-response items, with the weights being proportional to the information supplied by the two classes of items. The judges' ratings, in both metrics, and their associated errors of measurement are shown in Table H-3.

Table H-3
Cutpoints for Achievement Levels

Level	Mean Percent Correct, Multiple-choice and Short Constructed-response (Round 3)	Mean Paper Rating, Extended Constructed-response (Round 3)	Scale Score*	Standard Error of Scale Score**
Grade 4				
Basic	38	2.72	212	2.5
Proficient	62	3.14	243	2.1
Advanced	80	3.48	275	8.8
Grade 8				
Basic	41	2.13	244	2.6
Proficient	66	2.66	283	0.8
Advanced	85	3.22	328	7.7
Grade 12				
Basic	41	2.42	269	7.9
Proficient	67	2.85	304	2.8
Advanced	86	3.14	348	4.1

---

*Scale score is derived from a weighted average of the mean percents correct for multiple-choice and short constructed-response items and the mean paper ratings for extended constructed-response items after both were mapped onto the NAEP scale.

**The standard error of the scale is estimated from the difference in mean scale scores for the two equivalent subgroups of judges.

In the final stage of the mapping process, the achievement level cut points on the IRT measurement scale were combined over content areas and rescaled to the NAEP score scale. Weighted averages of the achievement level cut points were computed. The weighting constants accounted for the measurement precision of the test items evaluated by the panelists, the proportion of items belonging to each NAEP content area, and the linear NAEP scale transformations. These weighted averages produced the final cut points for the basic, proficient, and advanced achievement levels within each grade.

Figure H-1

Final Descriptions of 1992 Reading Achievement Levels

PREAMBLE

Reading for meaning involves a dynamic, complex interaction between and among the reader, the text, and the context. Readers, for example, bring to the process their prior knowledge about the topic, their reasons for reading it, their individual reading skills and strategies, and their understanding of differences in text structures.

The texts used in the reading assessment are representative of common real world reading demands. Students at grade 4 are asked to respond to literary and informational texts which differ in structure, organization, and features. Literary texts include short stories, poems, and plays that engage the reader in a variety of ways, not the least of which is reading for fun. Informational texts include selections from textbooks, magazines, encyclopedias, and other written sources whose purpose is to increase the reader's knowledge.

In addition to literary and informational texts, students at grades 8 and 12 are asked to respond to practical texts (e.g., bus schedules or directions for building a model airplane) that describe how to perform a task.

The context of the reading situation includes the purposes for reading that the reader might use in building a meaning of the text. For example, in reading for literary experience, students may want to see how the author explores or uncovers experiences, or they may be looking for vicarious experience through the story's characters. On the other hand, the student's purpose in reading informational texts may be to learn about a topic (such as the Civil War or the oceans) or to accomplish a task (such as getting somewhere, completing a form, or building something).

The assessment asks students at all three grades to build, extend, and examine text meaning from four stances or orientations:

> **Initial Understanding**—Students are asked to provide the overall or general meaning of the selection. This includes summaries, main points, or themes.
>
> **Developing Interpretation**—Students are asked to extend the ideas in the text by making inferences and connections. This includes making connections between cause and effect, analyzing the motives of characters, and drawing conclusions.
>
> **Personal Response**—Students are asked to make explicit connections between the ideas in the text and their own background knowledge and experiences. This includes comparing story characters with themselves or people they know, for example, or indicating whether they found a passage useful or interesting.
>
> **Critical Stance**—Students are asked to consider how the author crafted a text. This includes identifying stylistic devices such as mood and tone.

Figure H-1 (continued)

Final Descriptions of 1992 Reading Achievement Levels

These stances are not considered hierarchical or completely independent of each other. Rather, they provide a frame for generating questions and considering student performance at all levels. All students at all levels should be able to respond to reading selections from all of these orientations. What varies with students' developmental and achievement levels is the amount of prompting or support needed for response, the complexity of the texts to which they can respond, and the sophistication of their answers.

INTRODUCTION

The following achievement-level descriptions focus on the interaction of the reader, the text, and the context. They provide some specific examples of reading behaviors that should be familiar to most readers of this document. The specific examples are not inclusive; their purpose is to help clarify and differentiate what readers performing at each achievement level should be able to do. While a number of other reading achievement indicators exist at every level, space and efficiency preclude an exhaustive listing.

It should also be noted that the achievement levels are cumulative from basic to proficient to advanced. One level builds on the previous levels such that knowledge at the proficient level presumes mastery of the basic level, and knowledge at the advanced level presumes mastery at both the basic and proficient.

Grade 4—Basic

Fourth-grade students performing at the **basic level** *should demonstrate an understanding of the overall meaning of what they read.* When reading texts appropriate for fourth graders, *they should be able to make relatively obvious connections between the text and their own experiences.*

For example, when reading **literary text**, they should be able to tell what the story is generally about—providing details to support their understanding—and be able to connect aspects of the stories to their own experiences.

When reading **informational text**, basic-level fourth graders should be able to tell what the selection is generally about or identify the purpose for reading it; provide details to support their understanding; and connect ideas from the text to their background knowledge and experiences.

Grade 4—Proficient

Fourth grade students performing at the **proficient level** *should be able to demonstrate an overall understanding of the text, providing inferential as well as literal information.* When reading text appropriate to fourth grade, *they should be able to extend the ideas in the text by making inferences,*

Figure H-1 (continued)

Final Descriptions of 1992 Reading Achievement Levels

drawing conclusions, and making connections to their own experiences. The connection between the text and what the student infers should be clear.

For example, when reading **literary text**, proficient-level fourth graders should be able to summarize the story, draw conclusions about the characters or plot, and recognize relationships such as cause and effect.

When reading **informational text**, proficient-level students should be able to summarize the information and identify the author's intent or purpose. They should be able to draw reasonable conclusions from the text, recognize relationships such as cause and effect or similarities and differences, and identify the meaning of the selection's key concepts.

Grade 4—Advanced

Fourth-grade students performing at the **advanced level** *should be able to generalize about topics in the reading selection and demonstrate an awareness of how authors compose and use literary devices.* When reading text appropriate to fourth grade, *they should be able to judge texts critically and, in general, give thorough answers that indicate careful thought.*

For example, when reading **literary text**, advanced-level students should be able to make generalizations about the point of the story and extend its meaning by integrating personal experiences and other readings with the ideas suggested by the text. They should be able to identify literary devices such as figurative language.

When reading **informational text**, advanced-level fourth graders should be able to explain the author's intent by using supporting material from the text. They should be able to make critical judgments of the form and content of the text and explain their judgments clearly.

Grade 8—Basic

Eighth-grade students performing at the **basic level** *should demonstrate a literal understanding of what they read and be able to make some interpretations.* When reading text appropriate to eighth grade, *they should be able to identify specific aspects of the text that reflect the overall meaning, recognize and relate interpretations and connections among ideas in the text to personal experience, and draw conclusions based on the text.*

For example, when reading **literary text**, basic-level eighth graders should be able to identify themes and make inferences and logical predictions about aspects such as plot and characters.

When reading **informative text**, they should be able to identify the main idea and the author's purpose. They should make inferences and draw conclusions supported by information in the text.

Figure H-1 (continued)

Final Descriptions of 1992 Reading Achievement Levels

They should recognize the relationships among the facts, ideas, events, and concepts of the text (e.g., cause and effect and chronological order).

When reading **practical text**, they should be able to identify the main purpose and make predictions about the relatively obvious outcomes of procedures in the text.

Grade 8—Proficient

Eighth-grade students performing at the **proficient level** *should be able to show an overall understanding of the text, including inferential as well as literal information.* When reading text appropriate to eighth grade, *they should extend the ideas in the text by making clear inferences from it, by drawing conclusions, and by making connections to their own experiences—including other reading experiences.* Proficient eighth graders *should be able to identify some of the devices authors use in composing text.*

For example, when reading **literary text,** students at the proficient level should be able to give details and examples to support themes that they identify. They should be able to use implied as well as explicit information in articulating themes; to interpret the actions, behaviors, and motives of characters; and to identify the use of literary devices such as personification and foreshadowing.

When reading **informative text**, they should be able to summarize the text using explicit and implied information and support conclusions with inferences based on the text.

When reading **practical text**, proficient-level students should be able to describe its purpose and support their views with examples and details. They should be able to judge the importance of certain steps and procedures.

Grade 8—Advanced

Eighth-grade students performing at the **advanced level** *should be able to describe the more abstract themes and ideas of the overall text.* When reading text appropriate to eighth grade, they *should be able to analyze both meaning and form and support their analyses explicitly with examples from the text; they should be able to extend text information by relating it to their experiences and to world events.* At this level, student *responses should be thorough, thoughtful, and extensive.*

For example, when reading **literary text**, advanced-level eighth graders should be able to make complex, abstract summaries and theme statements. They should be able to describe the interactions of various literary elements (i.e., setting, plot, characters, and theme); to explain how the use of literary devices affects both the meaning of the text and their response to the author's style. They should be able critically to analyze and evaluate the composition of the text.

Figure H-1 (continued)

Final Descriptions of 1992 Reading Achievement Levels

When reading **informative text**, they should be able to analyze the author's purpose and point of view. They should be able to use cultural and historical background information to develop perspectives on the text and be able to apply text information to broad issues and world situations.

When reading **practical text**, advanced-level students should be able to synthesize information that will guide their performance, apply text information to new situations, and critique the usefulness of the form and content.

Grade 12—Basic

Twelfth-grade students performing at the **basic level** *should be able to demonstrate an overall understanding and make some interpretations of the text.* When reading text appropriate to twelfth grade, they *should be able to identify and relate aspects of the text to its overall meaning, recognize interpretations, make connections among and relate ideas in the text to their personal experiences, and draw conclusions.* They *should be able to identify elements of an author's style.*

For example, when reading **literary text**, twelfth-grade students should be able to explain the theme, support their conclusions with information from the text, and make connections between aspects of the text and their own experiences.

When reading **informational text**, basic-level twelfth graders should be able to explain the main idea or purpose of a selection and use text information to support a conclusion or make a point. They should be able to make logical connections between the ideas in the text and their own background knowledge.

When reading **practical text**, they should be able to explain its purpose and the significance of specific details or steps.

Grade 12—Proficient

Twelfth-grade students performing at the **proficient level** *should be able to show an overall understanding of the text, which includes inferential as well as literal information.* When reading text appropriate to twelfth grade, they *should be able to extend the ideas of the text by making inferences, drawing conclusions, and making connections to their own personal experiences and other readings. Connections between inferences and the text should be clear, even when implicit.* These *students should be able to analyze the author's use of literary devices.*

When reading **literary text**, proficient-level twelfth graders should be able to integrate their personal experiences with ideas in the text to draw and support conclusions. They should be able to explain the author's use of literary devices such as irony or symbolism.

Figure H-1 (continued)

Final Descriptions of 1992 Reading Achievement Levels

When reading **informative text**, they should be able to apply text information appropriately to specific situations and integrate their background information with ideas in the text to draw and support conclusions.

When reading **practical texts**, they should be able to apply information or directions appropriately. They should be able to use personal experiences to evaluate the usefulness of text information.

Grade 12—Advanced

Twelfth-grade students performing at the **advanced level** *should be able to describe more abstract themes and ideas in the overall text.* When reading text appropriate to twelfth grade, they *should be able to analyze both the meaning and the form of the text and explicitly support their analyses with specific examples from the text.* They *should be able to extend the information from the text by relating it to their experiences and to the world.* Their *responses should be thorough, thoughtful, and extensive.*

For example, when reading **literary text**, advanced-level twelfth graders should be able to produce complex, abstract summaries and theme statements. They should be able to use cultural, historical, and personal information to develop and explain text perspectives and conclusions. They should be able to evaluate the text, applying knowledge gained from other texts.

When reading **informational text**, they should be able to analyze, synthesize, and evaluate points of view. They should be able to identify the relationship between the author's stance and elements of the text. They should be able to apply text information to new situations and to the process of forming new responses to problems or issues.

When reading **practical texts,** advanced-level twelfth graders should be able to make a critical evaluation of the usefulness of the text and apply directions from the text to new situations.

Figure H-2

Draft Descriptions of the Achievement Levels
Prepared by the Original Level-setting Panel

4th-Grade Draft Descriptions

BASIC performance in reading should include:

- Determining what a text is about
- Identifying characterizations, settings, conflicts, or plots in a story
- Supporting one's understanding of a text with appropriate details
- Explaining why one likes or dislikes a text
- Connecting material in a text to personal experiences
- Making predictions about situations beyond the confines of a text
- Demonstrating an ability to maintain a focus over the entirety of a longer text

PROFICIENT performance in reading should include:

- Summarizing a text
- Recognizing an author's intent or purpose
- Making simple inferences based on information provided in a text
- Using information from a text to draw a basic conclusion
- Determining the meaning of key concepts in the text and connecting them to the main idea
- Recognizing the progression of ideas and the cause-and-effect relationships in a text
- Using the surrounding text to assign meaning to a word or phrase

ADVANCED performance in reading should include:

- Explaining an author's intent, using supporting material from the text
- Describing the similarities and differences in characters
- Demonstrating an awareness of the use of literary devices and figurative language
- Applying inferences drawn from a text to personal experiences
- Extending the meaning of a text by integrating experiences and information outside of the text
- Making and explaining a critical judgment of a text
- Demonstrating an ability to adapt reading purpose to genre and/or writing style

8th-Grade Draft Descriptions

BASIC performance in reading should include:

- Identifying the main idea or purpose of a text using information both stated and implied
- Expressing an author's purpose, viewpoint, and/or theme

Figure H-2 (continued)

Draft Descriptions of the Achievement Levels
Prepared by the Original Level-setting Panel

* Using information from a text to draw and support conclusions
* Making inferences appropriate to the information provided in a text
* Recognizing the cause-and-effect relationships in a text
* Making logical connections from the material in a text to personal knowledge and experience

PROFICIENT performance in reading should include:

* Restating the main idea using supportive details and examples from a text
* Summarizing a text using information both stated and implied
* Making inferences from a text in order to draw valid conclusions
* Interpreting the actions, behaviors, and motives of characters
* Integrating personal knowledge and experience to enhance one's understanding of a text
* Identifying an author's use of literary devices

ADVANCED performance in reading should include:

* Describing how specific literary elements interact with each other
* Synthesizing the information in a text to obtain abstract meaning or to perform a task
* Finding new applications for information derived from a text
* Making personal and critical evaluations of a text
* Analyzing an author's purpose, viewpoint, and/or theme
* Explaining an author's use of literary devices

12th-Grade Draft Descriptions

BASIC performance in reading should include:

* Explaining the main idea of a text
* Describing the main purpose in reading a selection
* Recognizing the significance of details from a reading in order to support a conclusion or perform a task
* Applying the information gathered from reading to meet an objective or support a conclusion
* Explaining the basic elements of an author's literary devices

PROFICIENT performance in reading should include:

* Drawing conclusions from and making inferences about information from different texts and writing styles
* Integrating background information with newly acquired information to support conclusions

Figure H-2 (continued)

Draft Descriptions of the Achievement Levels
Prepared by the Original Level-setting Panel

- Applying information from a text in an appropriate manner
- Bringing personal experience and accumulated knowledge into the process of critically evaluating a text
- Explaining an author's purpose in using complex literary devices

ADVANCED performance in reading should include:

- Providing innovative elaborations from textual information
- Analyzing and evaluating different points of view by means of comparison and contrast
- Identifying the relationships between an author's or narrator's stance and the various elements of the text
- Critically evaluating a text within a specific frame of reference
- Bringing the knowledge of other texts to the process of critical evaluation
- Using cultural or historical information provided in a text to develop perspectives on other situations
- Using cultural or historical information to develop perspectives on a text

Figure H-3

Revised Draft Descriptions of the Achievement Levels Recommended by the Follow-up Validation Panel

Revised 4th-Grade Draft Descriptions

Basic performance in reading should include:

- Determining what a story/informational text is about (i.e. topic, main idea)
- Determining the main purpose for reading a selection
- Identifying character(s), setting(s), conflict(s), or plot(s) in a story
- Supporting one's understanding of a story/informational text with appropriate details
- Explaining why one likes or dislikes what they have read [a reading]
- Connecting material from a story/informational text to personal experiences
- Making predictions about situations beyond the confines of the printed material
- Maintaining a focus over the entirety of a story/informational text

Proficient performance in reading should include:

- Summarizing a story/informational text
- Recognizing an author's intent or purpose
- Making simple inferences based on information provided in a story/informational text
- Drawing a valid conclusion from a story/informational text
- Determining the meaning of key concepts in the story/informational text and connecting them to the main idea
- Recognizing relationships in a story/informational text (time order, cause/effect, compare/contrast)

Advanced performance in reading should include:

- Explaining an author's intent, using supporting material from the story/informational text
- Describing the similarities and difference in characters, settings, and plots
- Demonstrating an awareness of the use of literary devices, such as figurative language
- Applying inferences drawn from a story/informational text to personal experiences
- Extending the meaning of a story/informational text by integrating experiences and information outside of the text
- Making and explaining a critical judgment of a story/informational text
- Demonstrating an ability to adapt reading purpose to a variety of printed material and/or writing style

Revised 8th-Grade Draft Descriptions

Basic performance in reading should include:

- Identifying the main idea, theme, or purpose of a text
- Describing the main purpose for reading a selection

Figure H-3 (continued)

Revised Draft Descriptions of the Achievement Levels
Recommended by the Follow-up Validation Panel

- Expressing an author's purpose and viewpoint
- Making inferences, predictions, and drawing conclusions that are supported by information in a text
- Recognizing the relationships among facts, ideas, events, and concepts within a text (e.g., cause and effect, chronological order, and characterization)
- Making logical connections between the text and personal knowledge
- Maintaining a focus over the entirety of a story/informational text

Proficient performance in reading should include:

- Restating the main idea, theme, or purpose of a text using supporting details and examples
- Summarizing a text using both stated and implied information
- Interpreting the actions, behaviors, and motives of characters
- Using personal knowledge and experience to enhance one's understanding of a text
- Identifying an author's use of literary devices (i.e. personification, foreshadowing, and so forth).
- Using inferences from a text in order to draw valid conclusions.

Advanced performance in reading should include:

- Describing how specific literary elements (i.e., setting, plot, characters, and theme) interact with each other
- Synthesizing the information in a text to obtain implied meaning or to perform a task
- Applying information derived from a text to new situations.
- Explaining an author's use of literary devices (i.e., irony, personification, and foreshadowing)
- Responding personally and critically to a text
- Analyzing an author's purpose and viewpoint
- Using cultural or historical information to develop perspectives on a text
- Using cultural or historical information provided in a text to develop perspectives on other situations

Revised 12th-Grade Draft Descriptions

Basic performance in reading should include:

- Explaining the main idea, theme, or purpose of a text
- Describing the main purpose for reading a selection
- Recognizing the significance of details from a reading in order to support a conclusion or perform a task

Figure H-3 (continued)

Revised Draft Descriptions of the Achievement Levels
Recommended by the Follow-up Validation Panel

- Applying the information gathered from reading to meet an objective or support a conclusion
- Identifying and explaining the basic elements of an author's literary devices
- Making logical connections between a text and personal knowledge and experience
- Maintaining a focus over the entirety of a story/informational text

Proficient performance in reading should include::

- Drawing conclusions and making inferences from different texts and writing styles
- Integrating background information with newly acquired information to support conclusions
- Applying information from a text in an appropriate manner
- Applying personal experience and accumulated knowledge to the process of critically evaluating a text
- Explaining an author's purpose in using complex literary devices (i.e. irony, symbolism)

Advanced performance in reading should include:

- All basic and proficient reading behaviors listed previously
- Prompted by information from a text, innovating in new situations and creating new answers to old situations
- Analyzing, synthesizing, and evaluating different points of view by means of comparison and contrast
- Identifying the relationships between an author's or narrator's stance and the various elements of the text
- Critically evaluating a text within a frame of reference
- Applying the knowledge of other texts to the process of critical evaluation
- Using cultural or historical information to develop perspectives on a text
- Using cultural or historical information provided in a text to develop perspectives on other situations

Figure H-4

Meeting Participants, NAEP Reading Achievement Level Setting
Original Meeting, St. Louis, Missouri, August 21 - 25, 1992

David Awbrey
Wichita Eagle
Wichita, KS

Dorothy Botham
Milwaukee Public Library
Milwaukee, WI

Anna Caballero
Attorney
Salinas, CA

Kathy Casseday
WFSP Radio Station
Kingwood, WV

Dee Ellis
Trimble Banner Newspaper
Milton, KY

Nona Smith
NAACP
New York, NY

Lillaine Speese
Oakdale Elementary School
Oroville, CA

Clifton Whetten
Retired Construction Sprvsr.
Elfrida, AZ

P. Richard Brackett
Brackett & Assoc. Motivational Marketing
Company
Brentwood, TN

Kathleen Harkey
Corporate Presentations
Nashville, TN

Patricia Oliverez
Salinas Public Library
Salinas, CA

Christine Sentz
North Milwaukee Branch Library
Milwaukee, WI

Carolyn Sullivan
Planters & Merchants Bank
Gillett, AR

Paula Abrams
City Hall
Bedford, KY

Rhonda Cantrell Dunn
Nashville Urban League
Nashville, TN

Harlon Gaskill (CPA)
Gaskill, Pharis & Pharis
Dalhart, TX

Jean McManis
Local/State Education Volunteer
State College, PA

Linda Borsum
Lakeview School District
Battlecreek, MI

Anne Kraut
Elementary Supervisor
Princeton, WV

Robert Williams
Macomb Intermediate SD
Clinton Township MI

Figure H-4 (continued)

Meeting Participants, NAEP Reading Achievement Level Setting Original Meeting, St. Louis, Missouri, August 21 - 25, 1992

Constance Boyd
Owen J. Roberts SD
King of Prussia, PA

Mary Gonzalez
Mesa Public Schools
Mesa, AZ

James Schindler
Jordan SD
Salt Lake City, UT

Kathryn Flannery
Indiana University
Bloomington, IN

Catherine Hatala
School District of Philadelphia
Philadelphia, PA

Raymond Morgan
Old Dominion University
Virginia Beach, VA

Berton Wiser
Columbus Public School
Columbus, OH

Freda Andrews
Durham Public Schools
Durham, NC

Tim Barnes
Ashdown Public Schools
Ashdown, AR

Larry Barretto
Maplewood Elementary School
Coral Springs, FL

Gloria Darling
Conway Public Schools
Conway, AR

Nina Frederick
Marion County School System
Hackleburg, AL

Karen Fugita
Oak Grove SD
San Jose, CA

Anne Gregory
Durham Public Schools
Durham, NC

Joseph Howard
Josiah Quincy School
West Roxbury, MA

Roberta Johnson
Cleveland Public Schools
Cleveland, OH

Marcia Jolicoeur
Lisbon Falls School
Lewiston, ME

Elizabeth Litchfield
Westwood School District
Emerson, NJ

Jean Young
Houston ISD
Houston, TX

Wilma Centers
Wolfe County Middle School
Campton, KY

Eunice Coakley
Greenville School
Greenville, SC

Eugenia Constantinou
Prince Georges County Schools
Silver Spring, MD

Figure H-4 (continued)

Meeting Participants, NAEP Reading Achievement Level Setting
Original Meeting, St. Louis, Missouri, August 21 - 25, 1992

Walt Cottingham
Henderson City Schools
Zirconia, NC

Deborah Davidson
Westhampton Beach UFSD
Patchogue, NY

Julia Dominique
Department of Education USVI
Sunnyisle, VI

Patricia Gerdes
Waelder ISD
Schulenburg, TX

Leslie Leech
Elkton School
Elkton, SD

Belva Leffel
Whittier Christian Jr. High
Norwalk, CA

Harriett McAllaster
Volusia County Schools
DeLand, FL

Mary Orear
Camden-Rockport HS & MS
Rockport, ME

Judith Zinsser
Houston ISD
Houston, TX

Mary Ann Ledbetter
East Baton Rouge Parish School Board
Baton Rouge, LA

Cora Cummins
Conway Public Schools
Conway, AR

Stanley Fraundorf
Cuba City Public Schools
Cuba City, WI

Georgia Howard
Volusia County Schools
Holly Hill, FL

Roger Larsen
Campbell County SD
Gillette, WY

Judith Lusk
Norfield School District
Rockbury, VT

Donnie McQuinn
Wolfe County Board of Education
Pine Ridge, KY

Meredith Powers
Swansea School
Providence, RI

Beth Schieber
Kingfisher Schools
Okarche, OK

Carolyn Sue Wilson
Greenville, SC

Sue Zak
Cleveland Board of Education
Garfield Heights, OH

Figure H-5

Meeting Participants, NAEP Reading Achievement Level Setting Follow-Up Validation Meeting, San Diego, California, October 9 - 11, 1992

Meredith Powers
Swansea School
Providence, RI

Roger Larsen
Campbell County SD
Gillett, WY

Beth Schieber
Kingfisher Schools
Okarche, OK

Elizabeth Litchfield
Westwood School District
Emmerson, NJ

Larry Barretto
Maplewood Elementary School
Coral Springs, FL

Anne Gregory
Durham Public Schools
Durham, NC

Debra Davidson
Westhampton Beach UFSD
Patchogue, NY

Eugenia Constantinou
Prince Georges County School
Silver Spring, MD

Eunice Coakley
Greenville School
Greenville, SC

Nancy Livingston
Brigham Young University
Salt Lake City, UT

Susan McIntyre
University Wisconsin-Eau Claire
Eau Claire, WI

Clyde Colwell
Norfolk Public School
Norfolk, VA

Jo Prather
Mississippi Department of Education
Jackson, MS

Mary Orear
Camden-Rockport HS & MS
Rockport, ME

Shelia Potter
Michigan Department of Education
Lansing, MI

Gene Jongsma
IRA Subcommittee Member
San Antonio, TX

Peggy Dutcher
Michigan Education Assessment Program
Lansing, MI

Martha Carter
Milwaukee Public Schools
Milwaukee, WI

Mark Conley
Michigan State University
Holt, MI

APPENDIX I

THE NAEP SCALE ANCHORING PROCESS FOR THE 1992 MATHEMATICS ASSESSMENT

The NAEP Scale Anchoring Process for the 1992 Mathematics Assessment

Ina V.S. Mullis and Eugene G. Johnson

Educational Testing Service

Introduction

Beginning with the 1984 assessments, NAEP has generally reported students' subject area proficiency on 0-to-500 scales. These scales are used to report achievement for students at the various grades or ages assessed, including differences between performance from assessment to assessment for the nation and for various subpopulations of interest. To date, NAEP has used item response theory techniques to develop proficiency scales for reading, mathematics, science, U.S. history, and civics.

Although average proficiency is an efficient summary measure, some of the most interesting NAEP results are those based on performance differences for different points in the scale distributions. To provide an interpretation for both the average results (What does a 306 on the 0-to-500 scale actually mean?) and changes in performance distributions (What does it mean that fewer students are reaching level 250?), NAEP invented a scale anchoring process to describe the characteristics of student performance at various levels along the scales—typically, at levels 200, 250, 300, and 350. The descriptions of student performance are presented in the reports accompanied by the percentages of students performing at or above the various scale levels.

Scale anchoring is a way of attaching meaning to a scale. Traditionally, meaning has been attached to educational scales by norm-referencing, that is, by comparing students at a particular scale level to other students. In contrast, the NAEP scale anchoring is accomplished by describing what students at selected levels know and can do.

The mathematics composite scale was anchored in 1990. Since the 1992 mathematics scales were linked to the 1990 scales and shared a common framework and a pool of common items with the 1990 assessment, it was expected that the 1990 anchor descriptions would still be appropriate for 1992. However, the anchoring process was conducted again on the 1992 data to update the descriptions to reflect the newer item types included in the assessment, and to permit the selection of items to exemplify each of the levels. Consequently, the anchoring of the 1992 mathematics composite was viewed as an enhancement of the 1990 anchoring, and, as anticipated, the 1992 descriptions are very similar to the 1990 descriptions, with some variations.

In brief, NAEP's scale anchoring procedure for the 1992 mathematics assessment—like the 1990 assessment—was based on comparing item-level performance by students at four levels on the 0-to-500 overall mathematics proficiency scale—levels 200, 250, 300, and 350. These values (corresponding to standard deviation units of 50 from the overall mean in 1990 of 250) are far enough apart to be noticeably different but not so far apart as to be trivial. This analysis delineated four sets of anchor items that discriminated between adjacent performance levels on the scale. The four sets of empirically derived anchor items were studied by a panel of distinguished mathematics educators, who carefully considered and articulated the types of knowledge, skills, and reasoning abilities demonstrated by correct responses to the items in each set. The 16 panelists and the NAEP staff involved in the process worked first in two independent groups to develop descriptions using the 1990 descriptions and the 1992 anchor items. As might be expected, the two sets of descriptions were quite similar, but not identical. Thus, the panelists subsequently met as a whole to review both sets of descriptions and decide how best to present the combined view of the entire group. Anchoring results for the 1992 mathematics assessment are presented in several different reports. Each report provides the descriptions accompanied by some or all (depending on the report) of the anchor items available for public release. For each grade level at which the item was administered, each item is accompanied by its overall proportion correct (overall p-value) for the total population assessed and the p-values for each anchor level. The various steps in the procedure are detailed below.

The Scale Anchoring Analysis

NAEP's scale anchoring is grounded in an empirical process whereby the scaled assessment results are analyzed to delineate sets of items that discriminate between adjacent performance levels on the scale[1]. For the 1992 mathematics assessment, as in the 1990 assessment, the levels were 200, 250, 300, and 350. For these four levels, items were identified that were likely to be answered correctly by students performing at a particular level on the scale and much less likely to be answered correctly by students performing at the next lower level.

To provide a sufficient pool of respondents, in identifying anchor items, students at level 200 were defined as those whose estimated mathematics proficiency (as defined by their first composite plausible value) was between 187.5 and 212.5, students at 250 were defined as those with estimated proficiency between 237.5 and 262.5, those at 300 had estimated proficiencies between 287.5 and 312.5, and those at 350 between 337.5 and 362.5. In theory, proficiency levels above 350 or below 200 could have been defined; however, so few students in the assessment performed at the extreme ends of the scale that it was not possible to do so.

The 1992 mathematics scale anchoring analysis was based on the scaled composite proficiency results for fourth, eighth, and twelfth graders participating in the 1992 national assessment. As illustrated below, for each item in the NAEP assessment, ETS determined the weighted percentage and raw frequency (unweighted count) for students at each of the four scale levels correctly answering the item. This was done for each of the grade levels at which the item was administered, and for the grade levels combined, if the item was administered at more than one grade level. For example, regardless of the grade level, the data for each item were analyzed as shown in the following sample.

---

[1] A detailed discussion of the theoretical underpinnings of scale anchoring can be found in Beaton and Allen (1992).

Sample Scale Anchoring Results				
Scale point	200	250	300	350
Weighted p-value	0.49	0.85	0.96	0.98
Raw frequency	902	1555	1271	276

It should be noted that the percentages of students answering the item correctly at each of the four scale levels differ from the proportion of students scoring above each score level and from the overall p-value for the total sample at any one grade level.

Because the extended constructed-response items were scored on an ordered scale with 5 scoring levels (0 to 4), the above procedure, which relies on the notion of a correct or an incorrect response to an item, had to be generalized. To fit into the anchoring framework, each extended constructed response item was converted into 4 pseudoitems by dichotomization at each of the values 1 through 4. Thus, the first pseudoitem was coded 1 for scores 1, 2, 3, and 4, and coded 0 otherwise; the second pseudoitem was coded 1 for scores 2, 3, 4 and coded 0 otherwise; the third pseudoitem was coded 1 for scores 3 and 4 and coded 0 otherwise; and the fourth pseudoitem was coded 1 for a score of 4 and coded 0 otherwise. These pseudoitems were then analyzed in the same manner as the items scored correct/incorrect.

As described below, criteria were applied to the scale-level results and an analysis conducted to delineate the items that discriminated between scale levels. Because it was the lowest level being defined, level 200 did not have to be analyzed in terms of the next lower level, but only for the percentage of students at that level answering the item correctly. More specifically, for an item to anchor at level 200:

1) The p-value for students at level 200 had to be greater than or equal to 0.65, and

2) the calculation of the p-value at that level had to have been based on at least 100 students to ensure adequate stability of the estimate of the p-value.

As an example, the following results are for an item anchoring at level 200:

Level 200 Anchor Item Results				
Scale point	200	250	300	350
Weighted p-value	0.65	0.89	0.98	1.00
Raw frequency	116	706	510	23

For an item to anchor at the remaining levels, additional criteria had to be met. For example, to anchor at level 250:

1) The p-value for students at level 250 had to be greater than or equal to 0.65;

2) the p-value for students at level 200 had to be less than or equal to 0.50;

3) the difference between the two p-values had to be at least 0.30; and

4) the calculations of the p-values at both levels 200 and 250 had to have been based on at least 100 students.

The following data set illustrates the results for a level 250 anchor item:

Level 250 Anchor Item Results

Scale point	200	250	300	350
Weighted p-value	0.38	0.75	0.89	0.98
Raw frequency	247	569	509	83

The principles used for level 250 were also used to identify anchor items at levels 300 and 350. For example, the following results were obtained for an item anchoring at level 300:

Level 300 Anchor Item Results

Scale point	200	250	300	350
Weighted p-value	0.11	0.28	0.83	1.00
Raw frequency	134	670	512	52

The results below are for an item anchoring at level 350:

Level 350 Anchor Item Results

Scale point	200	250	300	350
Weighted p-value	0.00	0.22	0.37	0.94
Raw frequency	50	324	585	241

In summary, for any given anchor item, 1) students at the item's anchor level are likely to answer the item correctly ($p_1 \geq .65$); students at the next lower level are somewhat unlikely to answer the item correctly ($p_2 \leq .50$); and students at the next lower level are less likely than students at the anchor level to answer the item correctly ($p_1 - p_2 \geq .30$). Collectively, as identified through this procedure, the 1992 NAEP mathematics items at each anchor level represented advances in students' understandings from one level to the next—mathematical areas where students at that level were more likely to answer items correctly than were students at the next lower level.

Preparing for the Mathematics Item Anchoring Panel Meeting

The analysis procedures described above yielded 22 items that anchored at level 200, 45 items at level 250, 59 items at level 300, and 43 items at level 350. Additionally, to provide some information for cross-referencing purposes, items that "almost anchored" were also identified. While these items did not satisfy the anchoring criteria, they did satisfy the following relaxed criteria: The p-value for students at the anchor level was at least .60, the p-value at the next lower level was no more than .55, the difference between the two p-values was at least .27, and the calculations of the p-values at both levels was based on at least 20 students. This procedure yielded additional items at each score point (level 200—8 items, level 250—27 items, level 300—29 items, level 350—34 items) that could be used for further context in developing descriptions. Of the 432 items included in the process, 149 (34%) anchored and 98 (23%) almost anchored. Table I-1 provides a breakdown of the number of anchored and almost anchored items by content area and by grade.

In preparation for use by the scale anchoring panelists, the items were placed in notebooks by section in the following order: anchored at 200, almost anchored at 200, anchored at 250, almost anchored at 250, anchored at 300, etc. Again, for further cross-referencing purposes, the remaining items in the assessment were also included in the notebook under the "did not anchor" heading. Each item was accompanied by its scoring guide (for constructed response items) and by the full anchoring documentation; that is, the anchoring information for each grade level at which an item was administered, the anchoring information across grades, the p-value for the total population of respondents at each grade level, and the mathematics content-area and process classifications for the items.

As described in *Mathematics Objectives, 1990 Assessment* (NAEP, 1988), which was also the framework for the 1992 assessment, the mathematics assessment was designed to measure five content areas, each with three ability levels. To ensure that the anchoring performance descriptions tied back to the assessment specifications, within anchor level sections, the items in the notebooks were sorted by the five content areas—Numbers and Operations; Measurement; Geometry; Data Analysis, Statistics, and Probability; and Algebra and Functions. Within content area, the items were sorted by ability level—procedural knowledge, conceptual understanding, and problem solving.

The Scale Anchoring Panel

Sixteen mathematics educators were invited to participate in the anchoring process. They represented teachers at the various grade levels involved, state mathematics supervisors from several of the states participating in the Trial State Assessment (including Washington, DC), large-city mathematics curriculum coordinators, and college mathematics professors and researchers. The group was also balanced by region of the country, race/ethnicity, and gender. (See Figure I-1 for a list of the participants.)

Table I-1
Counts of Items Anchoring and Almost Anchoring by Content Area and Grade

Content Area	DNA[*]	200	200A[**]	250	250A	300	300A	350	350A
GRADE 4:									
Num & Op	17	10	4	18	9	6	5	0	0
Measurement	6	7	1	5	6	4	2	0	0
Geometry	8	3	2	5	4	5	3	0	0
Data Analysis	9	0	0	10	1	2	1	0	0
Algebra & Fcns	8	1	0	6	0	1	2	0	0
GRADE 8:									
Num & Op	21	3	3	10	4	12	5	0	6
Measurement	6	1	2	4	4	7	1	6	4
Geometry	10	1	1	3	3	11	3	2	5
Data Analysis	16	0	0	3	4	7	3	1	3
Algebra & Fcns	8	2	0	3	1	9	2	2	5
GRADE 12:									
Num & Op	16	1	1	4	1	7	2	8	3
Measurement	11	1	0	1	0	3	1	8	3
Geometry	17	0	1	0	1	6	7	8	1
Data Analysis	21	0	0	1	1	7	0	2	0
Algebra & Fcns	29	1	0	0	1	5	1	17	5
ACROSS GRADES:									
Num & Op	41	10	4	18	11	14	10	8	8
Measurement	19	7	2	6	7	9	3	8	7
Geometry	32	3	2	5	4	14	9	8	6
Data Analysis	35	0	0	10	4	11	3	2	3
Algebra & Fcns	38	2	0	6	1	11	4	17	10

[*] Did not anchor
[**] Almost anchored

Figure I-1
Mathematics Scale Anchoring Panel

Charles Allen	Michigan Department of Education Lansing, Michigan
Thomas Carpenter	University of Wisconsin Madison, Wisconsin
John Dossey	Illinois State University Normal, Illinois
Edward Dubinsky	Purdue University West Lafayette, Indiana
David Farber	Voorhees High School Glen Gardener, New Jersey
Deborah Haimo	University of Missouri St. Louis, Missouri
Linda Kolnowski	Detroit Public Schools Detroit, Michigan
Gordon Lewis	Washington, DC, Public Schools Washington, DC
Mary Lindquist	Columbus College Columbus, Georgia
Donna Long	Indiana Department of Education Indianapolis, Indiana
Wendell Meeks	Illinois State Department of Education Springfield, Illinois
Tej Pandey	California Department of Education Sacramento, California
Mattye Pollard-Cole	Colorado State Department of Education Denver, Colorado
Diane Shaffer	Rhode Island Department of Education Providence, Rhode Island
Jocelyn Walton	Plainfield High School Plainfield, New Jersey
Charles Watson	Arkansas Department of Education Little Rock, Arkansas

The Process for Developing the Descriptions

The three-day anchoring meeting began on the morning of the first day, during which time panelists were thoroughly briefed in the anchoring process and given their assignment: With the objectives for the 1992 mathematics assessment and the anchor descriptions for the 1990 assessment as a reference, use the information in the anchor item notebooks to describe the mathematical knowledge, understandings, and problem-solving abilities demonstrated by the students at each anchor level in each of the five content areas. Based on the items anchoring at each anchor level (cross-referenced with "almost anchored" and "did not anchor" items), the panelists were asked to draft a description of achievement at each level in one-half page or less.

The meeting was structured so that the remainder of the first day and the entire second day could be devoted to the panelists working with staff in two independent groups to accomplish this task. In each of the independent groups, panelists and staff worked together to analyze the knowledge, skills, and reasoning abilities required by each item. Lists were developed portraying these for each mathematics content area at each anchor level. Based on these question-by-question analyses, which were prominently displayed around the room on poster paper, each group of panelists then drafted a description of performance for each anchor level. The two sets of draft descriptions appear as Figure I-2.

On the third day, the panelists and staff met as a whole to combine the two independently derived sets of descriptions. They also worked on developing short "titles" or descriptors for each category, and selecting example items to accompany the anchor level descriptions.

Both groups agreed that the two drafts were very similar and that with some final review and editing either set would have appropriately described the anchor item information. However, they did like the benefit of the cross-validation process and the fact that more people were able to participate in the process. As the group worked through the two descriptions, they identified preferences for some parts of each of the descriptions, resolved some issues, and made some formatting decisions. The combined view was checked by staff against the anchoring data, edited, and sent to the panelists for final review. The final draft of the descriptions is presented in Figure I-3.

Reporting the Anchor Item Results

Anchoring results are presented in three reports: the *NAEP 1992 Mathematics Report Card for the Nation and the States*, *Interpreting NAEP Scales*, and the *Data Compendium from the NAEP 1992 Mathematics Assessment for the Nation and the States*. In the first two reports, the anchor descriptions are supported by several items anchoring at each level. The *Data Compendium* includes all anchor items and almost anchor items available for public release. Each anchor item in the reports is, for each grade level at which it was assessed, accompanied by the overall percentage of success on the item as well as the anchor level information for each grade at which it was assessed. This is designed to reduce confusion between the percentages of success on the individual anchor items illustrating particular levels on the scale and the percentage of students who perform at or above each scale level.

Figure I-2
Draft Descriptions Prepared Independently by the Two Groups of Panelists

GROUP A

Level 200

Students at this level have a basic understanding of the whole numbers and their operations. They are able to write number names and select the largest four-digit number from a list. They can add and subtract most whole numbers without a calculator. They can add, subtract, multiply, and divide whole numbers with a calculator. In situations involving money they can round a number to the nearest dollar. These students can identify the solution to a one-step word problem.

These students can select appropriate instruments to measure length, weight, and temperature and can identify appropriate units to measure length. In geometry, these students can identify common shapes in two- and three-dimensions as well as select the results of flipping and turning shapes in the plane. They have a very elementary understanding of symmetry. In algebra, they are capable of extending a simple sequential symbol pattern.

Level 250

Students at this level have a solid understanding of whole number operations which allows them to translate between situations and mathematical representations for those settings. They can solve one- and simple two-step problems involving whole numbers, including interpreting the meaning of the resulting number. Their sense of whole numbers and their use extends to knowing when to estimate and what information in a problem situation may be extraneous.

GROUP B

Level 200

The students at this level function in the domain of whole numbers. They can add and subtract whole numbers, and when a calculator is available, they can multiply and divide. They can select the largest whole number from a set of numbers in the thousands, and can match the verbal and symbolic names for a number.

Length and weight are familiar attributes to these students in that they can select appropriate instruments and units to measure these attributes. They can recognize some basic properties of two-dimensional geometric figures as well as the names of standard examples of these figures. They can extend simple visual patterns.

Level 250

When presented with a problem situation, students have some understanding of the problem, they can identify extraneous information and have some knowledge of when to use computational estimation.

Students at this level have an understanding of addition, subtraction, multiplication, and division with whole numbers. They can solve routine one-step multiplication and division problems involving remainders and two-step addition and subtraction problems. They can round whole numbers

Figure I-2 (continued)
Draft Descriptions Prepared Independently by the Two Groups of Panelists

GROUP A

In measurement they can read scales on instruments, use a ruler to measure length in centimeters, and perform simple conversions of units in a system. They can use measurement scales in the solution of elementary word problems. Students at this level have extended abilities to deal with common planar shapes, including seeing them embedded in other figures or using them to dissect other figures. This emerging spatial sense allows them to visualize a cube and select counter-examples to elementary generalizations about the properties of a figure.

In data, these students are able to read, construct, and interpret data represented in tables, bar graphs, circle graphs and pictographs, including one- and two-step problems based on such data. They have an elementary understanding of relative frequency probability and related simple expectancy settings.

In the algebra dimension students at this level are able to read a number line and extend values to intermediate points. Their understanding also allows the extensions of simple arithmetic progression patterns in an applied setting.

Level 300

Students at this level are able to answer simple questions or solve simple problems involving fractions and decimals. They are able to both identify and create examples that illustrate equivalences of fractions and decimals, including locating the positions of such numbers on a number line. They can solve increasingly complex multi-step problems. Their understanding of percent includes the ability to calculate the effect of a

GROUP B

and solve simple word problems involving place value, estimation, and multiples.

These students can use a ruler to measure length and have some understanding of area and perimeter. They can solve simple problems using readings from instruments. They demonstrate a knowledge of properties of triangles, squares, rectangles, circles, and cubes. They can solve problems that require visualizing, drawing or manipulating simple geometric shapes. They can complete bar graphs and pictographs, as well as use information from graphs or tables to solve simple problems. They can recognize simple number patterns, are beginning to deal informally with the idea of a variable, and have some knowledge of simple probability.

Level 300

Students at this level can use various strategies and explain their reasoning in a variety of problem solving situations. They are able to solve problems involving not only whole numbers but with decimals and fractions. They can represent, compare, and find equivalent fraction, and use these concepts in solving routine problems. They can find a percent of a number and use this skill in simple problems. Multiplication and division of whole

Figure I-2 (continued)
Draft Descriptions Prepared Independently by the Two Groups of Panelists

GROUP A

percent increase on a total amount. With integers they are able to find simple products. At this level there is the emergence of some understanding of the number theory ideas of multiple and divisor.

In measurement these students can find areas of squares and rectangles and know the relationship between the perimeter of a square and the length of its side. They are able to give the measurement of the length of an object, use rulers with some flexibility, and use the scale on a map to approximate distance. Given a formula, they can substitute measures to get a numerical value.

In geometry, students at this level have a basic understanding of the properties of squares, rectangles, and parallelograms and are able to use those properties to identify necessary conditions and make some elementary indirect measurements. They are capable of finding the length of a missing side of a triangle in a simple similarity setting. They know that the sum of the measures of the angles of a triangle equal 180° and are able to use this property in simple problem settings. Using manipulatives they can combine shapes to represent a specified shape condition. Their spatial sense has increased to include the ability to visualize a cube in either three-space or its planar net arrangement.

In data analysis these students can draw data from a table to make decisions or, given additional data, insert new data in an existing table. They are able to draw data from circle and line graphs and compute with that data to answer questions or describe when an event occurred. They also have an understanding of bias in a sample. Their knowledge of probability is still rooted in relative frequencies in simple simulation situations. Students at this level can list the elements in a simple sample

GROUP B

and rational numbers have developed to the extent that students can use all four operations in multistep problems.

At this level student can read and use instruments in more complex situations. They can find areas of rectangles, recognize relationships among common units of measure, and use proportional relationships to solve routine problems involving similar triangles and scale drawings. In geometry, they have knowledge of definitions and properties of simple geometric figures in the plane. They can visualize composition and decomposition of two- and three-dimensional figures.

These students can calculate averages, select and interpret data from a variety of graphs, list the possible arrangements in a sample space, find the probability of a simple event, and have a beginning understanding of sample bias. They can evaluate algebraic expressions and solve linear inequalities by substitution, and solve equations involving square roots.

Figure I-2 (continued)
Draft Descriptions Prepared Independently by the Two Groups of Panelists

GROUP A

space and list the permutations of three objects.

In algebra these students can graph points on coordinate axes, locate the missing coordinates for a corner of a square, and identify which ordered pairs satisfy a given linear equation. With inequalities they can shade the points on a number line satisfying a simple interval given as an inequality, and solve elementary linear inequalities involving whole numbers. Students at this level show the ability to evaluate simple expressions and solve linear literal equations.

Level 350

Students at this level have extended their knowledge of number and variable to relate and represent large numbers using exponents and scientific notation. They can use exponents in evaluating the decimal value of a number displayed in scientific notation on a calculator or in solving equations involving powers of numbers. Their knowledge of percent extends to evaluating situations involving variables and estimating percents. Other situations reflect their ability to use rates in solving two-step application problems. In number theory they show a solid understanding of even and odd numbers and their properties under computation.

In measurement these students can solve a variety of perimeter and area problems involving triangles, quadrilaterals, and circles. Their concept of surface area allows the solution of problems involving rectangular solids. In geometry they can solve for the length of missing segments in more complex similarity situations involving the use of basic geometric theorems. Their overall knowledge of planar geometric concepts and relations has

GROUP B

Level 350

In problem solving, students are able to use their reasoning and analytic abilities when they encounter new situations. They can use reasoning strategies, data, models, and relevant mathematics in solving problems. Students can judge the reasonableness and correctness of their solutions.

Students at this level can reason and estimate with percents. They can recognize scientific notation and find the decimal equivalent. They can apply their knowledge of area and perimeter of simple geometric figures to solve problems. They can find the circumferences of circles and the surface areas of solid figures. Students can apply the Pythagorean theorem to fund the hypotenuse of a right triangle. They are beginning to use rectangular coordinates in problem-solving situations and can apply geometric properties and relationships in solving problems.

Figure I-2 (continued)
Draft Descriptions Prepared Independently by the Two Groups of Panelists

GROUP A

extended to the coordinate plane and includes slope, distance, and some ideas of the rate of change in linear settings. Students at this level know and are able to apply the Pythagorean theorem in a variety of settings.

In data these students can interpret information supplied by a graph of a step function and calculate the mean (average) from a table of grouped data. In combinatorial problems these students can list the possible occurrences and examine them to solve problems.

In algebra and functions, besides the growth of coordinate geometry, these students have an extended understanding of an ability to use the properties of exponents in equation solving and computation. They can solve complex literal equations and systems of linear equations. With functions, they can evaluate a quadratic function for a given value, as well as find the value for a composite function. Graphically they can identify the zeros of a function and the graphical effect of taking the absolute value of a given function. Their knowledge of trigonometry includes the ability to find the trigonometric value associated with an angle in a right triangle, evaluate a functional value of an angle given in radian measure on the unit circle, and identify the value of a trigonometric expression using a basic trigonometric identity. Additionally these students show the ability to evaluate and represent complex patterns involving both numbers and expressions including variables.

GROUP B

The students can compute means from frequency tables and create a sample space to determine probabilities. Students can use exponents and evaluate expressions given in functional notation. They can identify an equation describing a linear relation provided in a table, solve literal equations and systems of two linear equations. They have some beginning knowledge of trigonometric relations. They can interpret a graph to determine the zeros of a function, read values in a step function, and transform a graph by applying the absolute value. They are able to recognize patterns in order to solve problems.

1013

1014

Figure I-3
Description of Mathematics Proficiency for Four Anchor Levels on the NAEP Scale

Level 200	**Addition and Subtraction, and Simple Problem Solving with Whole Numbers**

Students at this level can identify solutions to one-step word problems, involving addition or subtraction. They can add and subtract whole numbers in most situations, and when a calculator is available, they can multiply and divide. They are able to select the largest whole number from a set of numbers in the thousands, and can match the verbal and symbolic names for numbers.

Students demonstrate familiarity with length and weight, by selecting appropriate instruments and units to measure these attributes. They are able to recognize some basic properties of two-dimensional geometric figures as well as the names of standard examples of these figures. They can recognize simple patterns.

Level 250	**Multiplication and Division, Simple Measurement, and Two-Step Problem Solving**

When presented with a problem situation, students at this level have some understanding of the problem, can identify extraneous information, and have some knowledge of when to use computational estimation. They have an understanding of addition, subtraction, multiplication, and division with whole numbers. They can solve simple two-step problems involving whole numbers. They are able to round whole numbers and solve simple word problems involving place value, estimation, and multiples.

Students can use a ruler to measure length in centimeters and have some understanding of area and perimeter. They can solve simple problems using readings from instruments. They demonstrate a knowledge of properties of triangles, squares, rectangles, circles, and cubes. They can solve problems that require visualizing, drawing or manipulating simple geometric shapes. They are able to complete bar graphs and pictographs, as well as use information from graphs or tables to solve simple problems. They can recognize simple number patterns, are beginning to deal informally with the idea of a variable, and have some knowledge of simple probability.

(continued)

Figure I-3 (continued)
Description of Mathematics Proficiency for Four Anchor Levels on the NAEP Scale

Level 300	**Reasoning and Problem-Solving Involving Fractions, Decimals, Percents, and Elementary Concepts in Geometry, Statistics, and Algebra**

Students at this level can use various strategies and explain their reasoning in a variety of problem-solving situations. They are able to solve problems involving not only whole numbers but with decimals and fractions. They can represent and find equivalent fractions, and use these concepts in solving routine problems. They can find a percent of a number and use this skill in simple problems. Multiplication and division of whole numbers have developed to the extent that students can use all four operations in multistep problems.

Students can read and use instruments in more complex situations. They can find areas of rectangles, recognize relationships among common units of measure, and solve routine problems involving similar triangles and scale drawings. They have knowledge of definitions and properties of simple geometric figures in the plane. Their spatial sense includes the ability to visualize a cube in either three-space or its flattened form in a plane.

Students can calculate averages, select and interpret data from a variety of graphs, list the possible arrangements in a sample space, find the probability of a simple event, and have a beginning understanding of sample bias. They can use knowledge of relative frequencies in simple simulation situations. Students show the ability to evaluate simple expressions and solve linear equations. Students can graph points on coordinate axes, locate the missing coordinates for a corner of a square, and identify which ordered pairs satisfy a given linear equation.

Level 350	**Reasoning and Problem Solving Involving Geometric Relationships, Algebra, and Functions**

Students at this level can reason and estimate with percents. They can recognize scientific notation and find the decimal equivalent. They can apply their knowledge of area and perimeter of simple geometric figures to solve problems. They can find the circumferences of circles and the surface areas of solid figures. They can solve for the length of missing segments in more complex similarity situations. Students can apply the Pythagorean theorem to find the hypotenuse of a right triangle. They are beginning to use rectangular coordinates in problem-solving situations and can apply geometric properties and relationships in solving problems.

Students can compute means from frequency tables and create a sample space to determine probabilities, and read the graph of a step function. Students can use exponents and evaluate expressions given in functional notation. In number theory, they have an understanding of even and odd numbers and their properties. They can identify an equation describing a linear relation provided in a table, and solve literal equations and systems of two linear equations. They have some knowledge of trigonometric relations. These students can represent and interpret complex patterns and data using numbers, expressions, and graphs. Given the graph of a function they can identify its zeros and the effect on the graph of taking the absolute value of the function.

APPENDIX J

THE NAEP SCALE ANCHORING PROCESS FOR THE 1992 READING ASSESSMENT

APPENDIX J

The NAEP Scale Anchoring Process for the 1992 Reading Assessment

Eugene G. Johnson, Ina V.S. Mullis, Jay R. Campbell, and Steven P. Isham

Educational Testing Service

Introduction

Beginning with the 1984 assessments, NAEP has generally reported students' subject area proficiency on 0-to-500 scales. These scales are used to report achievement for students at the various grades or ages assessed, including differences between performance from assessment to assessment for the nation and for various subpopulations of interest. To date, NAEP has used item response theory techniques to develop proficiency scales for reading, mathematics, science, writing, U.S. history, and civics.

Although average proficiency is an efficient summary measure, some of the most interesting NAEP results are those based on performance differences for different points in the scale distributions. To provide an interpretation for both the average results (What does a 306 on the 0-to-500 scale actually mean?) and changes in performance distributions (What does it mean that fewer students are reaching level 250?), NAEP invented a scale anchoring process to describe the characteristics of student performance at various levels along the scales—typically, at levels 200, 250, 300, and 350. The descriptions of student performance are presented in the reports accompanied by the percentages of students performing at or above the various scale levels.

Scale anchoring is a way of attaching meaning to a scale. Traditionally, meaning has been attached to educational scales by norm-referencing, that is, by comparing students at a particular scale level to other students. In contrast, the NAEP scale anchoring is accomplished by describing what students at selected levels know and can do.

On February 15-17, 1993, ETS applied a modified anchoring procedure to the 1992 reading achievement levels. As applied to the achievement levels, the anchoring process was designed to determine the sets of questions that students scoring at or above each achievement level cutpoint could perform with a high degree of success. A committee of reading experts, educators, and others was assembled to review the questions and, using their knowledge of reading and student performance, to generalize from the questions to descriptions of the types of skills exhibited at each achievement level.

The Scale Anchoring Analysis

A question was identified as anchoring at an achievement level for a given grade if it was answered correctly by at least 65 percent of the students in that grade scoring at the cutpoint of that achievement level, and by less than 65 percent of the students scoring at the cutpoints for any lower achievement level. In order to maximize the number of questions offered for consideration, the traditional discrimination criterion, which required that the chances of success at the next lower level be at least 30 percentage points lower, was not used.

To provide a sufficient pool of respondents in identifying anchor items, students at the cutpoint of each achievement level were defined as those whose estimated reading proficiency (as defined by their first composite plausible value) was within 12.5 points of the achievement level cutpoint on the NAEP scale. (The derivation of achievement level cutpoints on the NAEP scale is described in Appendix H.) This is consistent with previous anchoring procedures and provides an empirical estimate of the performance of students scoring at the cutpoint. To provide stable estimates, the calculations of the chances of success on an item had to be based on at least 75 students in the cutpoint interval; this was reduced from the previous requirement of 100 students to accommodate the small number of students reaching the advanced level.

The 1992 reading scale anchoring analysis was based on the scaled composite proficiency results for fourth, eighth, and twelfth graders participating in the 1992 national assessment. As illustrated below, for each item in the NAEP assessment, ETS determined the weighted percentage and raw frequency for students at each of the achievement levels correctly answering the item. This was done separately for each of the grade levels at which the item was administered. For example, the data for each item were analyzed as shown in the following sample.

Sample Scale Anchoring Results

Achievement level	Basic	Proficient	Advanced
Weighted p-value	0.22	0.49	0.73
Raw frequency	282	386	93

It should be noted that the percentages of students answering the item correctly at each of the achievement levels differ from the proportion of students scoring above each achievement level and from the overall p-value for the total sample at any one grade level.

Because the extended constructed response items were scored on an ordered scale with four scoring levels (minimal, partial, essential, and extensive), the above procedure, which relies on the notion of a correct or an incorrect response to an item, was generalized. To fit into the anchoring framework, each extended constructed-response item was treated as three distinct items corresponding to scores of partial or better, essential or better, and extensive. These distinct items were then analyzed in the same manner as items scored as correct/incorrect.

Thus, for example, an extended constructed-response item might anchor at the proficient level for partial or better responses, and at the advanced level for essential or better responses.

Because it was the lowest level being defined, the basic level did not have to be analyzed in terms of the next lower level, but only for the percentage of students at that level answering the item correctly. More specifically, for an item to anchor at the basic level:

1) The p-value for students at the basic level had to be greater than or equal to 0.65, and

2) the calculation of the p-value at that level had to have been based on at least 75 students to ensure adequate stability of the estimate of the p-value.

As an example, the following results are for an item anchoring at the basic level:

Basic Level Anchor Item Results

Achievement level	Basic	Proficient	Advanced
Weighted p-value	0.68	0.78	0.90
Raw frequency	308	413	115

For an item to anchor at the remaining levels, three criteria had to be met. For example, to anchor at the proficient level:

1) The p-value for students at the proficient level had to be greater than or equal to 0.65;

2) the p-value for students at the basic level had to be less than 0.65; and

3) the calculations of the p-values at both levels had to have been based on at least 75 students.

The following data set illustrates the results for a proficient level anchor item:

Proficient Level Anchor Item Results

Achievement level	Basic	Proficient	Advanced
Weighted p-value	0.34	0.73	0.95
Raw frequency	369	433	131

The same principles were used to identify anchor items at the advanced level. For example, the following results were obtained for an item anchoring at the advanced level:

Advanced Level Anchor Item Results			
Achievement level	Basic	Proficient	Advanced
Weighted p-value	0.13	0.41	0.84
Raw frequency	313	423	106

By anchoring the achievement level cutpoints, instead of the entire interval, it is possible to determine the types of skills exhibited by all students within an interval. Thus, an item anchoring at the basic level cutpoint will be answered correctly by at least 65 percent of minimally basic students and will be answered correctly by at least that percentage of students in the basic interval. Since the NAEP results are reported in terms of the percentages of students at or above each of the cutpoints, it is important to be able to say what all students in the interval are likely to be able to do. In contrast, an anchoring procedure based on the interval identifies skills that a typical member of the interval (e.g., a typical basic student) likely possesses. While we could infer what a typical student in the basic interval can likely do, we would not be able to infer the skills of a minimally basic student.

A description of an entire achievement level interval can be inferred by comparing the descriptions for adjacent cutpoints. Thus, the description for the basic cutpoint tells what all basic students are likely to be able to do with increasing certainty as their reading proficiency increases. The description of the proficient cutpoint refers to the abilities of minimally proficient students, but also provides information about the capabilities of basic students scoring at the top of the basic interval. To extend the description of the advanced achievement level, since that interval does not have an upper boundary, an additional set of questions were identified as "almost anchoring" at the advanced level. These questions had probabilities of success between 50 and 65 percent for minimally advanced students and identify the types of skills that more advanced students are likely to possess.

For example, the results below are for an item almost anchoring at the advanced level:

Almost Advanced Level Item Results			
Achievement level	Basic	Proficient	Advanced
Weighted p-value	0.11	0.31	0.55
Raw frequency	298	443	104

Preparing for the Reading Item Anchoring Panel Meeting

Table J-1 provides a breakdown of the numbers of anchored and almost anchored dichotomous items (i.e., items scored correct/incorrect) by content area and grade. The vast majority of these items anchored at some achievement level, or almost anchored at the advanced level. The remaining items that did not anchor were generally quite difficult.

Table J-2 provides similar information for the extended constructed response items that were scored on a four-point scale. As described above, each of these items was treated as three distinct items, corresponding to scores of partial or better, essential or better, and extensive. The counts in Table J-2 are in terms of these item parts. The item parts that did not anchor correspond to scores of extensive, and sometimes, essential or better.

In preparation for use by the scale anchoring panelists, the items were placed in notebooks by grade in the following order: anchored at basic, anchored at proficient, anchored at advanced, and almost anchored at advanced (chance of success between 50 and 65 percent at the advanced level). For cross-referencing purposes, the remaining items in the assessment were also included in the notebook under the "did not anchor" heading. (These were the items answered correctly by fewer than 50 percent of the students at the advanced cutpoint.) Each item was accompanied by its scoring guide (for constructed-response items), the chance of success on the item for students at each achievement level, the counts and weighted proportions of students at each level, the overall percent correct on the item for the total population of respondents, and the reading purpose and stance classifications for the item.

The anchoring process was further informed by results using the item mapping procedure. Item mapping provides additional information about the performance of students within each of the achievement level intervals, and of students who performed below the basic level. In item mapping, the items are arranged in the order of the proficiency level corresponding to a defined expected probability of success based on the item response theory parameters. The items, or short descriptions, are then displayed, along with the proficiency value associate with the selected probability of success. For consistency with the anchoring process, a .65 expected probability of success was used.

The Process for Developing the Descriptions

Twenty reading education experts participated in a three-day anchoring meeting. They represented teachers of the three grade levels, college professors, state curriculum supervisors, and researchers. (See Figure J-1 for a list of the participants.) The panelists were divided into three groups, one for each grade level. The grade-level groups worked independently for the most part, with periodic meetings across the three groups to reconcile views. With the framework for the 1992 reading assessment and the achievement level descriptions as a reference, panelists were asked to use the information in the anchor item notebooks and from the item mapping to describe the knowledge, skills, and reasoning abilities demonstrated by the students at the cutpoint of each achievement level. In addition, performance as depicted by the maps or items that almost anchored was taken as indicating beginning or emerging skills for students in the interval. Based on the items anchoring at each level and the item maps, the

Table J-1

Counts of Dichotomous Reading Items Anchoring by Content Area And Grade

Content Area	Basic	Proficient	Advanced	Almost Anchored at Advanced	Did Not Anchor
GRADE 4:					
Literary	15	12	9	2	1
Informational	13	10	10	1	4
GRADE 8:					
Literary	10	12	6	0	4
Informational	16	15	5	5	5
Task-oriented	8	17	3	2	3
GRADE 12:					
Literary	4	14	9	0	3
Informational	20	17	11	3	3
Task-oriented	17	8	6	2	3

Table J-2

Counts of Polytomous Item Parts[*] Anchoring by Content Area And Grade

Content Area	Basic	Proficient	Advanced	Almost Anchored at Advanced	Did Not Anchor
GRADE 4:					
Literary	1	1	2	2	5
Informational	2	1	2	2	5
GRADE 8:					
Literary	0	2	1	2	4
Informational	1	4	3	2	8
Task-oriented	2	0	2	1	3
GRADE 12:					
Literary	1	4	2	0	2
Informational	6	4	5	1	7
Task-oriented	1	1	3	0	3

---

[*] Each polytomous item was treated as three separate items corresponding to scores of partial or better, essential or better, and extensive.

panelists were asked to draft a description of achievement at each level. In drafting these descriptions, the panelists were instructed to consider the context of the assessment and to not overinfer skills from limited numbers of items. The draft descriptions were checked by staff against the anchoring data, edited, and sent to the panelists for final review. The final draft of the descriptions is presented in Figure J-2. Each achievement level at each grade corresponds to a cutpoint on the NAEP scale as described in Appendix H.

Figure J-1

Reading Scale Anchoring Panel

Eileen Baldwin
Trenton School District
Trenton, NJ

Margo Brill-Wigant
Department of Education
Santa Fe, NM

Miriam Chaplin
NCTE
Cherry Hill, NJ

Karen Costello
Connecticut State Dept. of Ed.
Hartford, CT

Eunice Greer
University of Illinois
Champagne, IL

Robert Harrison
English/Language Arts Coordinator
Charleston, WV

Diane Hoffman
Middle School Teacher
Sykesville, MD

Janet Jones
William D. Wade Elementary
Waldorf, MD

Barbara Kapinus
Council of Chief State School Officers
Washington, DC

Judith Langer
SUNY-Albany
Albany, NY

Patricia McConigal
5th Grade Teacher
Underhill, VT

Jim Martin-Rehrman
Westfield State College
Westfield, MA

Leslie Morrow-Mandel
Rutgers University
Scotch Plains, NJ

Susan Neuman
Temple University
Philadelphia, PA

Charles Peters
Oakland Schools
Waterford, MI

Gary Rice
Louisiana State University
Baton Rouge, LA

Timothy Shanahan
University of Illinois at Chicago
Chicago, IL

Robert Swartz
University of Massachusetts
Newtonville, MA

Gwendolyn Williams
K - 8 Reading Specialist
Landover, MD

Figure J-2

Anchor Descriptions of the Reading Achievement Levels

Grade 4 students ...

Basic (212)	**... understand uncomplicated narratives and high-interest informative texts, identify an obvious theme, locate explicit information, summarize parts of text, and evaluate characters' actions.**

Fourth-grade students at the basic level can read uncomplicated narratives with understanding. The *literary* texts at this level include fables and realistic fiction about familiar topics. These students can answer questions that focus on specific parts of the story. They are able to identify an obvious theme or message. They can take the perspective of characters that are familiar or similar to themselves and compare characters to each other. In addition, they can relate to the feelings of familiar characters, as well as interpret and evaluate the characters' actions.

Students at the basic level are able to gain information from high-interest *informative* texts. These students are successful when texts are structured as narratives and deal with relatively familiar topics. Students can search for and locate explicit information within the text, as well as provide evidence of straightforward comprehension of the text. They are able to select relevant information in order to provide a summarization focusing on part of the text. They can understand the sequence of events and identify situations described in the text. They can build simple inferences based on specific information. These students also are able to construct their own simple questions related to the passage.

Grade 4 students ...

Proficient (243)	**... understand and interpret less familiar texts, provide textual support for interpretations, generalize across text, identify relevant information, understand subtleties in aspects of a story, relate text to background experiences, and formulate simple questions.**

Fourth-grade students at the proficient level can form an understanding and extend the meaning of more difficult, unfamiliar *literary* pieces—those in culturally different or historical settings. They are able to respond to questions that require some interpretation. Some can construct responses to the story as a whole, as well as consider subtleties in aspects of the story. When given interpretations of the story, they can provide some justification and support for those interpretations. They are able to recognize multiple perspectives. In addition, they have the ability to connect information in the story to the author's purpose, as well as consider alternate possibilities for the story's development.

Students at the proficient level are able to gain information and to interpret the meaning of *informative* text that contains narrative elements and direct quotes. Their responses to increasingly more challenging questions provide evidence that they can search for, locate, select, prioritize, and apply relevant information. They can generalize across parts of the text. They

Figure J-2 (continued)

Anchor Descriptions of the Reading Achievement Levels

can relate information from the selection to their own background experiences and to inferences that are provided for them. They also are able to recognize an author's basic organizational pattern.

Grade 4 students ...

Advanced (275)	... interpret and examine meaning of text, summarize information across whole text, develop their own ideas about textual information, understand some literary devices, and begin to formulate more complex questions about text.

Fourth-grade students at the advanced level can form an understanding of what they read and extend, elaborate, and examine the meaning of *literary* texts. They can construct responses to a story by selecting relevant information and building their own interpretations that remain consistent with the text. They are able to summarize information across the whole story. They understand some literary devices, such as figurative language, and can interpret the author's intentions.

Students at the advanced level can gain information from what they read and can extend, elaborate, and examine the meaning of *informative* texts about less familiar topics. They are able to read for the purpose of gaining a more thorough understanding of a particular topic, and some can develop their own ideas based on the information presented in the passage. They can discriminate the relative importance of ideas in the text and are beginning to form more complex questions about the selection. They are able to provide an explanation of the author's techniques for presenting information.

Grade 8 students ...

Basic (244)	... understand familiar genres, recognize central theme or topic, identify the central purpose of practical documents, identify literal information, interpret and describe character traits, and connect information from across text.

Eighth-grade students' responses at the basic level demonstrate fundamental understandings of *literary* texts from familiar genres. These texts are not complex or abstract—they contain a single perspective and a central focus. These students can answer questions that focus on surface or literal understandings of the story. They can identify the basic theme of a story and can connect ideas within one section or across larger parts of the text. They are able to interpret and describe character traits.

Students' responses at the basic level demonstrate an ability to make concrete interpretations from *informative* texts (i.e., biographies, articles, informative narratives) that

Figure J-2 (continued)

Anchor Descriptions of the Reading Achievement Levels

present information in a relatively straightforward manner. These students can recognize the central purpose by interpreting information across a text and by using structural text features, such as subheadings, exemplification, and organizational patterns. They are able to locate and to recognize explicitly stated information, as well as to connect information in one section of text with that from other sections. They are able to recognize the reasons an author might include partial information.

Students at the basic level are able to locate guidelines or directions that are explicitly stated in practical *documents*. They demonstrate some familiarity with documents, as well as an understanding of their purpose and usefulness. They can connect information presented within one section of a text to information in another section. They can articulate a personal view or choice about a document and support their opinion. In addition, they can use explicit directions to produce a specific textual form or document type.

Grade 8 students ...

Proficient (283)	... move beyond surface understanding of a text or multiple texts, make inferences about characters and themes, link generalizations to specific details, support an opinion about text, recognize an author's intentions, and use a document to solve simple problems.

Eighth-grade students at the proficient level are able to move beyond surface understandings of *literary* texts (i.e., historical fiction, tales) to develop fuller interpretations. They can recognize and interpret overall messages or themes implied in a literary piece. They are able to connect and make inferences about essential elements of stories and characters. They are able to interpret a character's ideas and feelings based on the events in the story and their own interpretation of the character's personality and role. These students can develop a perspective on a character's motivation by relying on their own understanding of human nature and essential story features, such as plot, dialogue, and description. They also can recognize an author's intentions and identify an author's use of symbolism to convey a story theme.

Proficient readers are able to locate and integrate information from different sections of an *informative* text and across multiple texts. At this level, students are able to gain information from textbook chapters, as well as biographies, articles, and informative narratives. These students can recognize a generalization and link it to specific details within the text. They demonstrate the ability to compare and contrast, as well as summarize information from across the text. They are able to form personal opinions about the content and provide supportive examples from text. They demonstrate an ability to use knowledge of organizational structures to gain information.

Readers at the proficient level are able to use multiple sources (i.e., time tables, instructions, maps) to locate information explicitly stated in a *document*. They can interpret the meaning of graphic symbols, such as map legends. They show the ability to perform tasks that

Figure J-2 (continued)

Anchor Descriptions of the Reading Achievement Levels

involve extracting information embedded within a document. They are able to discriminate among similar sources in accessing information to perform a task and solve a simple problem. They can understand how and why authors use text features and the relationship among particular features within documents, such as illustrations and examples.

Grade 8 students ...

Advanced (328)	... compare and contrast information across multiple texts, connect inferences with themes, understand underlying meanings, integrate prior knowledge with text interpretations, and demonstrate some ability to evaluate the limitations of documents.

Eighth-grade students reading at the advanced level are able to extend *literary* interpretations by relating personal knowledge to story characters and events. They demonstrate an understanding of fairly abstract themes and provide personal reactions to overall themes. They are able to interpret underlying meanings and complexities of characterizations and plot developments. They are able to connect inferences about characters' motives and feelings with story themes and provide supporting evidence from the story. In addition, they can relate themes across genres and to real-world situations. They also demonstrate the ability to consider the author's use of literary devices and relate it to an underlying theme.

Advanced eighth-grade readers are able to understand, to interpret, and to evaluate information presented in *informative* text. They are able to compare and contrast information within a text and across multiple texts and various genres. They make use of illustrations to enhance their interpretations of text. They can locate specific information embedded within text. They draw on knowledge from other subject areas and take a historical perspective in developing interpretations about text information. These students demonstrate the ability to formulate opinions about the information they read and support their ideas with appropriate text-based evidence.

Eighth-grade students at the advanced level are able to locate and to use very specific, highly embedded information in a fairly complex *document*. They use multiple pieces of information from various locations within a document to complete a task or solve a real-world problem. Many are able to evaluate the presentation of information in a document, recognize its limitations, and suggest improvements.

Grade 12 students...

Basic (269)	... develop interpretations from a variety of texts, understand overall arguments, recognize explicit aspects of plot and characters, support global generalizations, respond personally to texts, use major document features to solve real-world problems.

Figure J-2 (continued)

Anchor Descriptions of the Reading Achievement Levels

Twelfth-grade students at the basic level can gain meaning and develop interpretations from a variety of *literary* works (i.e., first-person adventures, narrative poems, tales). They respond to literature in a straightforward manner and focus their interpretations on specific aspects of a story. They are able to recognize fairly explicit aspects of plot development and characterization. Students at this level demonstrate surface understanding of characters' motives and are able to understand and use dialogue in constructing meaning. They can focus their attention, gain meaning, and develop interpretations from a character's perspective as well as their own. They respond personally to particular portions of a piece and report their responses to textual evidence.

Students at the basic level are able to gain information and to understand specific issues as a result of reading a variety of *informative* texts (i.e., encyclopedia entries, journal accounts, textbook chapters, science periodicals, editorials, and biographical essays). Students can gain information from reading individual texts or multiple texts on the same topic. They are able to recognize general arguments and viewpoints. They can use information from across text segments to make and support global generalizations. They are able to recognize explicitly stated problems and their solutions, as well as important causal relationships. In addition, they demonstrate an understanding of the potential contribution of illustrations and captions to readers' comprehension and engagement. These students are able to evaluate the importance of a particular issue and formulate an opinion.

Twelfth-grade students reading at the basic level are able to respond to forms, schedules, and practical *documents* adhering to most directions or guidelines. Drawing on text clues, they recognize and are able to locate explicit information stated in a document. These students demonstrate an understanding of the use of labels to group ideas and mark sections within documents. They are able to infer the purpose for document guidelines and compare a task completed according to the guidelines with another related task. In addition, these students are able to use accompanying maps, legends, symbols, and timetables to solve real-world problems. Students at the basic level recognize the most obvious limitations of a document's applicability and present personal reactions in response to document information.

Grade 12 students...

Proficient (304)	...integrate background experiences and knowledge with meaning from a variety of texts, interpret characters' motives, consider differing points of view, interpret literary devices, identify text structure and writing style, and apply document information to solve complex problems.

Proficient readers are able to form interpretations and express overall responses to *literary* texts (i.e., first-person adventures, narrative poems, tales). Drawing on their personal knowledge, they can interpret characters' motives and feelings, perceive significant character traits, identify similarities between characters, and develop an understanding of evolving characterizations within a story. In addition, they are able to find textual evidence to support

1031

their assumptions about characters and their actions. By delving beneath surface language and events, proficient readers are able to develop an understanding of the underlying intentions and communicative intent of dialogue. These readers integrate personal experiences with narrative or poetic elements and bring their real-world perceptions of the human condition to their literary interpretations. They are able to interpret figurative language and the symbolism suggested by major story elements.

Proficient readers are able to gain and to interpret relevant information from an individual *informative* passage or across multiple passages (i.e., encyclopedia entries, journal accounts, textbook chapters, science periodicals, editorials, and biographical essays). They are able to consider differing points of view in developing an understanding of text. They recognize the contributions of various texts in gaining overall understanding of a particular topic and are able to evaluate the credibility of different sources. Proficient readers demonstrate familiarity with informative genres by identifying organizational forms and recognizing patterns in writing style used by the author. They also are able to draw on background knowledge to interpret textual information and determine text reliability. Their responses to this type of text demonstrate an ability to analyze and make judgements about informative material.

Readers at the proficient level demonstrate comprehension of moderately complex and specific instructions presented in practical *documents*, including forms and schedules. Their responses demonstrate a clear understanding of a document's purpose. They are able to search documents to locate specific information from major sections and highly embedded details. They exhibit strategies for extracting and applying document information in successfully completing a multistep task. These readers are able to suggest alternative approaches to task completion and make choices based on an appropriate interpretation of the document's main features. They are able to access and use tabular and graphic information in making generalizations and decisions about real-world problems. They understand the purpose of a particular document and are able to tell the importance of complying with the guidelines.

Grade 12 students...

Advanced (348)	... construct complex understandings of multiple genres, interpret multidimensional aspects of characters, connect discipline-specific knowledge to text, examine author's craft, judge the value of informative sources, and suggest improvements for documents.

Advanced students are able to construct more complex and abstract understandings of *literary* texts by integrating personal knowledge and experiences with textual ideas and events. They are able to connect ideas and to relate interpretations across multiple types of literary genres. They are able to interpret the significance of major story elements, as well as draw on underlying meaning to develop a thorough understanding of an abstract theme. They consider non-explicit implications of language and dialogue within a literary piece. Drawing on their knowledge of human nature, they are able to interpret and describe nuances and multidimensional aspects of character relationships, feelings, and motives. They demonstrate an ability to examine their own personal understandings based on considerations of text meaning and real-world issues. They make use of their familiarity with literary elements to develop in-depth interpretations and examine critically the author's style and use of literary devices.

Students reading at the advanced level demonstrate the ability to synthesize and critically examine information presented in individual and multiple *informative* texts. They use information presented within a text to build overall understandings of conditions occurring across time. These readers cad identify the significance of events and draw on general background experiences, as well as discipline-specific knowledge to advance their understanding of information presented within text. They use genre-appropriate strategies to glean specific information, search for evidence to support generalizations, evaluate the credibility of multiple sources and identify potentially different uses for information gained from different sources. They perceive ways in which a point of view is expressed in an author's language and make judgements about the author's intent. By considering a text's purpose, structure, and content they are able to make and support judgements about its informative value.

Advanced readers demonstrate an ability to manage various organizational structures in accessing and applying information presented in documents, including forms and schedules. They are able to use specified directions and guidelines to complete highly detailed tasks. In addition, they are able to integrate text with graphic organizers in interpreting the meaning of written directions. These students are able to follow a series of complex steps specified by document directions in order to extract relevant information for a particular purpose. Based on a thorough examination of document text and structure, they make thoughtful and appropriate recommendations for improving the usefulness and presentation of information within a document.

APPENDIX K

Constructed-response Item Score Statistics

Appendix K

CONSTRUCTED-RESPONSE ITEM SCORE STATISTICS

This appendix contains information about the constructed-response items included in the scaling of data from the 1992 main assessments of reading, mathematics, and writing.

The information in the tables includes, for each subject area and age/grade, the NAEP item numbers for each of the constructed-response items included in scaling, and the block that contains the item. The tables also indicate the codes from the NAEP database that denote the range of responses and the correct responses. A portion of the responses to the constructed-response items were scored twice for the purpose of examining rater reliability. For each item, the number of papers with responses that were scored a second time is listed, along with the percent agreement between raters and an index of reliability based on those responses. Cohen's Kappa (Cohen, 1968) is the reliability estimate used for dichotomized items. For items that are not dichotomized (i.e., polytomous items), the intraclass correlation coefficient is used as the index of reliability.

Table K-1
Score Range, Percent Agreement, and Cohen's Kappa*
for the Short Constructed-response Reading Items Used in Main Assessment Scaling
Age 9/Grade 4

Item	Block	Range of Response Codes	Correct Response Codes	Sample Size	Percent Agreement	Cohen's Kappa
R012002	RC	1-2	2	559	90	0.82
R012004	RC	1-2	2	559	87	0.76
R012008	RC	1-2	2	559	91	0.85
R012010	RC	1-2	2	559	89	0.82
R012102	RD	1-2	2	543	88	0.70
R012104	RD	1-2	2	543	93	0.87
R012106	RD	1-2	2	543	87	0.78
R012108	RD	1-2	2	543	93	0.86
R012109	RD	1-2	2	543	93	0.87
R012112	RD	1-2	2	543	91	0.85
R012601	RE	1-2	2	559	89	0.74
R012604	RE	1-2	2	559	90	0.79
R012611	RE	1-2	2	559	89	0.84
R012201	RF	1-2	2	550	84	0.69
R012206	RF	1-2	2	550	93	0.89
R012208	RF	1-2	2	550	88	0.82
R012210	RF	1-2	2	550	90	0.83
R012702	RG	1-2	2	534	89	0.76
R012703	RG	1-2	2	534	92	0.86
R012705	RG	1-2	2	534	90	0.80
R012706	RG	1-2	2	534	85	0.76
R012710	RG	1-2	2	534	92	0.88
R012303	RH	1-2	2	578	88	0.78
R012306	RH	1-2	2	578	83	0.72
R012308	RH	1-2	2	578	88	0.82
R012310	RH	1-2	2	578	91	0.87
R012403	RI	1-2	2	559	87	0.76
R012406	RI	1-2	2	559	82	0.72
R012407	RI	1-2	2	559	86	0.78
R012409	RI	1-2	2	559	86	0.78
R012503	RJ	1-2	2	570	92	0.85
R012504	RJ	1-2	2	570	92	0.88
R012506	RJ	1-2	2	570	92	0.87
R012508	RJ	1-2	2	570	95	0.93
R012511	RJ	1-2	2	570	92	0.87

---

* Cohen's Kappa is a measure of reliability that is appropriate for items that are dichotomized.

Table K-2
Score Range, Percent Agreement, and Intraclass Correlation for the Extended Constructed-response Reading Items Used in Main Assessment Scaling
Age 9/Grade 4

Item	Block	Range of Response Codes	Sample Size	Percent Agreement	Intraclass Correlation
R012006	RC	1-4	559	90	0.95
R012111	RD	1-4	543	95	0.98
R012607	RE	1-4	559	93	0.94
R012204	RF	1-4	550	91	0.96
R012708	RG	1-4	534	96	0.97
R012305	RH	1-4	578	88	0.89
R012401	RI	1-4	559	94	0.93
R012512	RJ	1-4	570	94	0.98

Table K-3
Score Range, Percent Agreement, and Cohen's Kappa*
for the Short Constructed-response Reading Items Used in Main Assessment Scaling
Age 13/Grade 8

Item	Block	Range of Response Codes	Correct Response Codes	Sample Size	Percent Agreement	Cohen's Kappa
R013102	RC	1-2	2	620	81	0.63
R013104	RC	1-2	2	620	85	0.73
R013105	RC	1-2	2	620	90	0.80
R013108	RC	1-2	2	620	96	0.85
R013110	RC	1-2	2	620	84	0.76
R013111	RC	1-2	2	620	88	0.78
R012803	RD	1-2	2	613	82	0.69
R012807	RD	1-2	2	613	94	0.91
R012810	RD	1-2	2	613	80	0.66
R012813	RD	1-2	2	613	87	0.80
R012601	RE	1-2	2	633	85	0.73
R012604	RE	1-2	2	633	84	0.71
R012611	RE	1-2	2	633	91	0.76
R013203	RF	1-2	2	592	95	0.79
R013205	RF	1-2	2	592	86	0.74
R013207	RF	1-2	2	592	84	0.70
R013209	RF	1-2	2	592	88	0.75
R013211	RF	1-2	2	592	84	0.76
R012702	RG	1-2	2	637	93	0.70
R012705	RG	1-2	2	637	83	0.70
R012706	RG	1-2	2	637	74	0.53
R012703	RG	1-2	2	637	82	0.67
R012710	RG	1-2	2	637	91	0.82
R012713	RG	1-2	2	637	96	0.92
R012901	RH	1-2	2	606	79	0.62
R012905	RH	1-2	2	606	85	0.77
R012907	RH	1-2	2	606	80	0.68
R012909	RH	1-2	2	606	90	0.82
R012910	RH	1-2	2	606	88	0.78
R012914	RH	1-2	2	606	85	0.76
R013302	RI	1-2	2	635	98	0.94
R013304	RI	1-2	2	635	94	0.84
R013307	RI	1-2	2	635	94	0.91
R013310	RI	1-2	2	635	83	0.68
R013402	RJ	1-2	2	607	97	0.94
R013405	RJ	1-2	2	607	85	0.74
R013407	RJ	1-2	2	607	92	0.86
R013409	RJ	1-2	2	607	84	0.73

---

* Cohen's Kappa is a measure of reliability that is appropriate for items that are dichotomized.

Table K-3 (continued)
Score Range, Percent Agreement, and Cohen's Kappa*
for the Short Constructed-response Reading Items Used in Main Assessment Scaling
Age 13/Grade 8

Item	Block	Range of Response Codes	Correct Response Codes	Sample Size	Percent Agreement	Cohen's Kappa
R013411	RJ	1-2	2	607	84	0.74
R013412	RJ	1-2	2	607	88	0.78
R013001	RK	1-2	2	619	81	0.59
R013003	RK	1-2	2	619	98	0.97
R013005	RK	1-2	2	619	88	0.76
R013007	RK	1-2	2	619	94	0.88
R013008	RK	1-2	2	619	86	0.78
R013009	RK	1-2	2	619	86	0.76
R013010	RK	1-2	2	619	88	0.78
R013011	RK	1-2	2	619	80	0.70
R014702	RM	1-2	2	595	86	0.74
R014703	RM	1-2	2	595	87	0.62
R014704	RM	1-2	2	595	85	0.74
R014706	RM	1-2	2	595	87	0.77
R014709	RM	1-2	2	595	87	0.76
R014710	RM	1-2	2	595	81	0.64

---

* Cohen's Kappa is a measure of reliability that is appropriate for items that are dichotomized.

Table K-4
Score Range, Percent Agreement, and Correlation for the Extended Constructed-response Reading Items Used in Main Assessment Scaling
Age 13/Grade 8

Item	Block	Range of Response Codes	Sample Size	Percent Agreement	Correlation Correlation
R013106	RC	1-4	619	89	0.93
R012808	RD	1-4	613	91	0.91
R012607	RE	1-4	633	90	0.94
R013201	RF	1-4	592	86	0.90
R013212	RF	1-4	592	94	0.96
R012708	RG	1-4	637	90	0.93
R012903	RH	1-4	605	87	0.94
R013312	RI	1-4	635	92	0.93
R013403	RJ	1-4	607	96	0.96
R013406	RJ	1-4	607	88	0.95
R013004	RK	1-4	619	88	0.94
R014705	RM	1-4	595	87	0.94
R014713	RM	1-4	595	92	0.96

Table K-5
Score Range, Percent Agreement, and Cohen's Kappa*
for the Short Constructed-response Reading Items Used in Main Assessment Scaling
Age 17/Grade 12

Item	Block	Range of Response Codes	Correct Response Codes	Sample Size	Percent Agreement	Cohen's Kappa
R013102	RC	1-2	2	539	79	0.60
R013104	RC	1-2	2	539	87	0.70
R013105	RC	1-2	2	539	87	0.76
R013108	RC	1-2	2	539	95	0.81
R013110	RC	1-2	2	539	83	0.74
R013111	RC	1-2	2	539	85	0.74
R013501	RD	1-2	2	555	88	0.77
R013503	RD	1-2	2	555	91	0.83
R013505	RD	1-2	2	555	91	0.81
R013508	RD	1-2	2	555	90	0.85
R013509	RD	1-2	2	555	92	0.88
R013602	RE	1-2	2	537	88	0.80
R013604	RE	1-2	2	537	86	0.65
R013605	RE	1-2	2	537	82	0.69
R013607	RE	1-2	2	537	91	0.83
R013609	RE	1-2	2	537	90	0.79
R013611	RE	1-2	2	537	90	0.85
R013203	RF	1-2	2	554	95	0.60
R013205	RF	1-2	2	554	88	0.71
R013207	RF	1-2	2	554	88	0.67
R013209	RF	1-2	2	554	84	0.71
R013211	RF	1-2	2	554	84	0.68
R013701	RG	1-2	2	561	86	0.48
R013702	RG	1-2	2	561	87	0.69
R013704	RG	1-2	2	561	85	0.73
R013708	RG	1-2	2	561	81	0.60
R013710	RG	1-2	2	561	85	0.76
R013712	RG	1-2	2	561	88	0.80
R013801	RH	1-2	2	524	79	0.59
R013803	RH	1-2	2	524	79	0.59
R013806	RH	1-2	2	524	77	0.58
R013808	RH	1-2	2	524	84	0.71
R013809	RH	1-2	2	524	82	0.71
R013810	RH	1-2	2	524	89	0.80
R013302	RI	1-2	2	543	98	0.96
R013304	RI	1-2	2	543	93	0.85
R013307	RI	1-2	2	543	91	0.85
R013310	RI	1-2	2	543	87	0.76

---

* Cohen's Kappa is a measure of reliability that is appropriate for items that are dichotomized.

Table K-5 (continued)
Score Range, Percent Agreement, and Cohen's Kappa*
for the Short Constructed-response Reading Items Used in Main Assessment Scaling
Age 17/Grade 12

Item	Block	Range of Response Codes	Correct Response Codes	Sample Size	Percent Agreement	Cohen's Kappa
R013402	RJ	1-2	2	521	98	0.96
R013405	RJ	1-2	2	521	87	0.67
R013407	RJ	1-2	2	521	92	0.82
R013409	RJ	1-2	2	521	88	0.74
R013411	RJ	1-2	2	521	82	0.71
R013412	RJ	1-2	2	521	91	0.80
R013902	RK	1-2	2	556	92	0.78
R013903	RK	1-2	2	557	91	0.84
R013904	RK	1-2	2	557	94	0.81
R013906	RK	1-2	2	557	79	0.65
R013908	RK	1-2	2	557	89	0.73
R013910	RK	1-2	2	557	92	0.86
R013913	RK	1-2	2	556	91	0.85
R015503	RM	1-2	2	541	92	0.68
R015505	RM	1-2	2	541	85	0.66
R015509	RM	1-2	2	541	89	0.81
R015512	RM	1-2	2	541	89	0.75
R015604	RN	1-2	2	522	92	0.83
R015607	RN	1-2	2	522	84	0.72

---

* Cohen's Kappa is a measure of reliability that is appropriate for items that are dichotomized.

Table K-6
Score Range, Percent Agreement, and Correlation
for the Extended Constructed-response Reading Items Used in Main Assessment Scaling
Age 17/Grade 12

Item	Block	Range of Response Codes	Sample Size	Percent Agreement	Intraclass Correlation
R013106	RC	1-4	539	85	0.92
R013506	RD	1-4	555	97	0.99
R013610	RE	1-4	537	96	0.98
R013201	RF	1-4	554	85	0.91
R013212	RF	1-4	554	92	0.95
R013706	RG	1-4	561	91	0.94
R013805	RH	1-4	524	91	0.97
R013312	RI	1-4	543	91	0.94
R013403	RJ	1-4	521	94	0.97
R013406	RJ	1-4	521	87	0.96
R013915	RK	1-4	556	87	0.93
R015507	RM	1-4	541	88	0.94
R015514	RM	1-4	541	93	0.98
R015601	RN	1-4	522	90	0.94
R015609	RN	1-4	522	90	0.96
R015612	RN	1-4	522	85	0.91

Table K-7
Score Range, Percent Agreement, and Cohen's Kappa* for the Short Constructed-response Mathematics Items Used in Main Assessment Scaling
Age 9/Grade 4

Item	Block	Range of Response Codes	Correct Response Codes	Sample Size	Percent Agreement	Cohen's Kappa
M039201	MC	1-4	4	576	99	0.97
M039301	MC	1-4	4	576	99	0.99
M040001	MC	1-5	5	576	98	0.96
M040201	MC	1-2	2	576	94	0.86
M022201	ME	1-3	2-3	589	92	0.86
M022501	ME	1-5	5	589	94	0.89
M022801	ME	1-3	3	589	98	0.97
M022802	ME	1-2	2	589	98	0.97
M019701	MF	1-2	2	567	99	0.98
M019801	MF	1-3	2-3	567	89	0.77
M019901	MF	1-3	2-3	567	98	0.97
M020001	MF	1-2	2	567	98	0.97
M020101	MF	1-2	2	567	96	0.91
M020201	MF	1-2	2	567	95	0.92
M020301	MF	1-4	4	567	99	0.98
M020401	MF	1-2	2	567	98	0.97
M020501	MF	1-2	2	567	96	0.93
N277903	NF	1-2	2	567	98	0.96
M020701	MF	1-4	4	567	91	0.86
M044601	MG	1-3	3	591	95	0.91
M045101	MG	1-4	3-4	591	98	0.96
M045301	MG	1-4	4	591	97	0.95
M010631	MH	1-3	3	572	98	0.95
M040301	MI	1-2	2	570	97	0.96
M040901	MI	1-3	2-3	570	98	0.96
M061901	MJ	1-3	3	563	93	0.89
M061902	MJ	1-3	2-3	563	97	0.95
M061903	MJ	1-2	2	563	97	0.95
M061904	MJ	1-3	2-3	563	97	0.92
M061905	MJ	1-4	4	563	94	0.88
M061906	MJ	1-3	3	563	97	0.91
M046001	MK	1-5	5	582	98	0.97
M046601	MK	1-4	4	582	98	0.96
M046801	MK	1-5	5	582	99	0.98
M046901	MK	1-5	5	582	99	0.99
M047301	MK	1-4	4	582	98	0.97
M041301	ML	1-5	5	600	99	0.97
M041701	ML	1-4	4	600	99	0.99

---

* Cohen's Kappa is a measure of reliability that is appropriate for items that are dichotomized.

Table K-7 (continued)
Score Range, Percent Agreement, and Cohen's Kappa*
for the Short Constructed-response Mathematics Items Used in Main Assessment Scaling
Age 9/Grade 4

Item	Block	Range of Response Codes	Correct Response Codes	Sample Size	Percent Agreement	Cohen's Kappa
M041901	ML	1-3	2-3	600	98	0.97
M042002	ML	1-4	4	600	99	0.99
M042003	ML	1-3	3	600	99	0.98
M042401	ML	1-3	3	600	99	0.98
M042501	ML	1-2	2	600	95	0.92
M043201	MM	1-4	4	574	98	0.97
M043301	MM	1-4	4	574	97	0.93
M043401	MM	1-4	4	574	98	0.97
M043402	MM	1-4	4	574	97	0.95
M043403	MM	1-3	3	574	97	0.92
M043601	MN	1-2	2	581	98	0.91
M043801	MN	1-4	4	581	99	0.98
M044202	MN	1-5	5	581	99	0.98
M048701	MO	1-2	2	597	96	0.92
M048801	MO	1-3	3	597	98	0.97
M048901	MO	1-4	3-4	597	98	0.93

---

* Cohen's Kappa is a measure of reliability that is appropriate for items that are dichotomized.

Table K-8

Score Range, Percent Agreement, and Correlation for the Extended Constructed-response Mathematics Items Used in Main Assessment Scaling

Age 9/Grade 4

Item	Block	Range of Response Codes	Sample Size	Percent Agreement	Intraclass Correlation
M045401	MG	1-4	591	84	0.87
M041201	MI	1-4	570	75	0.82
M043501	MM	1-4	574	89	0.94
M044401	MN	1-4	581	92	0.93
M049001	MO	1-4	597	78	0.86

Table K-9
Score Range, Percent Agreement, and Cohen's Kappa*
for the Short Constructed-response Mathematics Items Used in Main Assessment Scaling
Age 13/Grade 8

Item	Block	Range of Response Codes	Correct Response Codes	Sample Size	Percent Agreement	Cohen's Kappa
M050801	MC	1-2	2	589	99	0.99
M050901	MC	1-4	4	589	98	0.97
M051001	MC	1-2	2	589	93	0.89
M022201	ME	1-3	2-3	654	91	0.81
M022501	ME	1-5	5	654	95	0.90
M022801	ME	1-3	3	654	99	0.97
M022802	ME	1-2	2	654	99	0.98
M023701	ME	1-2	2	654	98	0.96
M019701	MF	1-2	2	633	99	0.97
M019801	MF	1-3	2-3	633	95	0.90
M019901	MF	1-3	2-3	633	99	0.96
M020001	MF	1-2	2	633	98	0.96
M020101	MF	1-2	2	633	99	0.98
M020201	MF	1-2	2	633	96	0.89
M020301	MF	1-4	4	633	100	0.99
M020401	MF	1-2	2	633	99	0.98
M020501	MF	1-2	2	633	98	0.97
M020801	MF	1-6	6	633	98	0.96
M020901	MF	1-2	2	633	90	0.85
M021001	MF	1-2	2	633	99	0.97
M021101	MF	1-3	3	633	92	0.87
M021201	MF	1-3	3	633	95	0.92
M021301	MF	1-2	2	633	96	0.94
M021302	MF	1-2	2	633	95	0.92
M044601	MG	1-3	3	635	97	0.93
M045101	MG	1-4	3-4	635	98	0.96
M045301	MG	1-4	4	635	97	0.95
M045601	MG	1-4	4	635	97	0.95
M045701	MG	1-4	4	635	97	0.95
M013031	MH	1-4	4	607	99	0.98
M013131	MH	1-3	3	607	96	0.92
M052401	MI	1-2	2	636	89	0.82
M052901	MI	1-3	2-3	636	86	0.74
M053001	MI	1-2	2	636	91	0.85
M061901	MJ	1-3	3	644	91	0.79
M061903	MJ	1-2	2	644	98	0.94
M061904	MJ	1-3	2-3	644	95	0.92
M061902	MJ	1-3	2-3	644	96	0.91

---

* Cohen's Kappa is a measure of reliability that is appropriate for items that are dichotomized.

Table K-9 (continued)
Score Range, Percent Agreement, and Cohen's Kappa*
for the Short Constructed-response Mathematics Items Used in Main Assessment Scaling
Age 13/Grade 8

Item	Block	Range of Response Codes	Correct Response Codes	Sample Size	Percent Agreement	Cohen's Kappa
M061907	MJ	1-3	3	644	93	0.82
M061908	MJ	1-3	3	644	96	0.83
M061905	MJ	1-4	4	644	93	0.88
M046001	MK	1-5	5	619	99	0.96
M046601	MK	1-4	4	619	98	0.96
M046801	MK	1-5	5	619	100	0.99
M046901	MK	1-5	5	619	100	0.99
M047301	MK	1-4	4	619	100	0.98
M047901	MK	1-3	3	619	99	0.97
M054001	ML	1-2	2	601	96	0.86
M054101	ML	1-2	2	601	98	0.97
M051201	MM	1-2	2	662	99	0.96
M051301	MM	1-3	3	662	99	0.97
M051601	MM	1-3	3	662	97	0.93
M052101	MM	1-6	5-6	662	99	0.98
M054801	MN	1-4	3-4	639	92	0.86
M055201	MN	1-5	4-5	639	98	0.96
M048701	MO	1-2	2	661	90	0.82
M048801	MO	1-3	3	661	99	0.98
M048901	MO	1-4	3-4	661	97	0.95
M049801	MO	1-2	2	661	94	0.87

---

* Cohen's Kappa is a measure of reliability that is appropriate for items that are dichotomized.

Table K-10
Score Range, Percent Agreement, and Correlation
for the Extended Constructed-response Mathematics Items Used in Main Assessment Scaling
Age 13/Grade 8

Item	Block	Range of Response Codes	Sample Size	Percent Agreement	Intraclass Correlation
M051101	MC	1-4	589	74	0.77
M045901	MG	1-4	635	83	0.82
M053101	MI	1-4	636	88	0.91
M054301	ML	1-4	601	88	0.87
M052201	MM	1-4	662	87	0.92
M055501	MN	1-4	639	83	0.88

Table K-11
Score Range, Percent Agreement, and Cohen's Kappa*
for the Short Constructed-response Mathematics Items Used in Main Assessment Scaling
Age 17/Grade 12

Item	Block	Range of Response Codes	Correct Response Codes	Sample Size	Percent Agreement	Cohen's Kappa
M056801	MC	1-2	2	575	98	0.94
M056901	MC	1-5	5	575	94	0.87
M057001	MC	1-5	5	575	98	0.96
M057101	MC	1-2	2	575	95	0.89
M023901	ME	1-3	3	552	99	0.98
M024701	ME	1-4	3-4	552	96	0.94
M025302	ME	1-3	3	552	91	0.82
M021401	MF	1-2	2	591	99	0.98
M021501	MF	1-2	2	591	97	0.94
M021502	MF	1-2	2	591	99	0.97
M021601	MF	1-4	4	591	94	0.80
M021602	MF	1-2	2	591	93	0.85
M020201	MF	1-2	2	591	98	0.89
M020301	MF	1-4	4	591	100	0.99
M020401	MF	1-2	2	591	98	0.96
M020501	MF	1-2	2	591	99	0.98
M020801	MF	1-6	6	591	99	0.97
M020901	MF	1-2	2	591	94	0.90
M021001	MF	1-2	2	591	99	0.99
M021101	MF	1-3	3	591	94	0.88
M021201	MF	1-3	3	591	95	0.91
M021701	MF	1-2	2	591	98	0.96
M021702	MF	1-2	2	591	96	0.94
M021801	MF	1-2	2	591	97	0.94
M057201	MG	1-2	2	610	99	0.97
M057501	MG	1-2	2	610	99	0.97
M058001	MG	1-2	2	610	97	0.96
M058201	MG	1-5	4-5	610	99	0.97
M058301	MG	1-3	3	610	98	0.95
M058401	MG	1-4	4	610	98	0.95
M013031	MH	1-4	4	551	98	0.97
M013131	MH	1-3	3	551	97	0.94
M011931	MH	1-2	2	551	97	0.93
M012031	MH	1-3	3	551	98	0.97
M052401	MI	1-2	2	611	89	0.80
M053301	MI	1-2	2	611	92	0.86
M061901	MJ	1-3	3	561	94	0.79
M061904	MJ	1-3	2-3	561	92	0.85

---

* Cohen's Kappa is a measure of reliability that is appropriate for items that are dichotomized.

Table K-11 (continued)
Score Range, Percent Agreement, and Cohen's Kappa*
for the Short Constructed-response Mathematics Items Used in Main Assessment Scaling
Age 17/Grade 12

Item	Block	Range of Response Codes	Correct Response Codes	Sample Size	Percent Agreement	Cohen's Kappa
M061907	MJ	1-3	3	561	93	0.86
M061908	MJ	1-3	3	561	94	0.81
M061905	MJ	1-4	4	561	91	0.81
M062101	MJ	1-3	3	561	98	0.97
M058901	MK	1-3	3	566	98	0.97
M059702	MK	1-4	4	566	96	0.92
M059801	MK	1-3	2-3	566	95	0.90
M054001	ML	1-2	2	596	96	0.92
M054401	ML	1-4	3-4	596	97	0.95
M054501	ML	1-4	3-4	596	97	0.90
M060001	MM	1-4	3-4	587	91	0.83
M060201	MM	1-4	3-4	587	98	0.96
M060401	MM	1-4	4	587	98	0.96
M060501	MM	1-5	4-5	587	95	0.91
M060601	MM	1-4	4	587	97	0.93
M054801	MN	1-4	3-4	564	90	0.80
M055201	MN	1-5	4-5	564	98	0.96
M055601	MN	1-4	4	564	98	0.95
M060801	MO	1-2	2	567	98	0.95
M061001	MO	1-4	4	567	99	0.98

---

* Cohen's Kappa is a measure of reliability that is appropriate for items that are dichotomized.

Table K-12
Score Range, Percent Agreement, and Correlation
for the Extended Constructed-response Mathematics Items Used in Main Assessment Scaling
Age 17/Grade 12

Item	Block	Range of Response Codes	Sample Size	Percent Agreement	Intraclass Correlation
M053401	MI	1-4	611	73	0.91
M062401	MJ	1-4	561	90	0.92
M054601	ML	1-4	596	92	0.86
M060701	MM	1-4	587	80	0.85
M055701	MN	1-4	564	93	0.86
M061801	MO	1-4	567	89	0.84

Table K-13
Score Range, Percent Agreement, and Correlation for the Constructed-response Writing Items Used in Main Assessment Scaling
Age 9/Grade 4

Item	Block	Range of Response Codes	Sample Size	Percent Agreement	Intraclass Correlation
W001001	WC	1-6	528	85	0.94
W001601	WD	1-6	559	85	0.93
W001101	WE	1-6	606	86	0.95
W001201	WF	1-6	562	84	0.93
W001701	WG	1-6	595	80	0.90
W001301	WH	1-6	543	82	0.92
W001401	WI	1-6	571	87	0.94
W001801	WJ	1-6	541	83	0.93
W001501	WK	1-6	565	83	0.92

Table K-14
Score Range, Percent Agreement, and Correlation for the Constructed-response Writing Items Used in Main Assessment Scaling
Age 13/Grade 8

Item	Block	Range of Response Codes	Sample Size	Percent Agreement	Intraclass Correlation
W002001	WC	1-6	625	85	0.94
W001601	WD	1-6	621	81	0.93
W002101	WE	1-6	595	76	0.87
W002201	WF	1-6	600	72	0.86
W001701	WG	1-6	585	73	0.86
W002301	WH	1-6	604	73	0.86
W002401	WI	1-6	597	91	0.94
W001801	WJ	1-6	575	89	0.91
W002501	WK	1-6	596	88	0.92
W002601	WL	1-6	648	80	0.91
W001901	WM	1-6	617	77	0.91

Table K-15
Score Range, Percent Agreement, and Correlation for the Constructed-response Writing Items Used in Main Assessment Scaling
Age 17/Grade 12

Item	Block	Range of Response Codes	Sample Size	Percent Agreement	Intraclass Correlation
W002001	WC	1-6	598	83	0.91
W002701	WD	1-6	544	92	0.97
W002101	WE	1-6	580	83	0.91
W002201	WF	1-6	536	70	0.84
W002801	WG	1-6	540	67	0.81
W002301	WH	1-6	539	73	0.86
W002401	WI	1-6	534	84	0.84
W002901	WJ	1-6	566	81	0.87
W002501	WK	1-6	534	83	0.81
W002601	WL	1-6	557	81	0.89
W003001	WM	1-6	547	71	0.81
W003101	WN	1-6	544	78	0.85

APPENDIX L

Differential Item Functioning (DIF) Results

Table L-1
Reading Items Identified as "C" Items in at least One Comparison*

Item	Block	Scale	Grade	Comparison	Group Favored
R012501	RJ	Information	4	White/Black	Black
R012511	RJ	Information	4	White/Black	White
R012912	RH	Information	8	White/Black	White
R013001	RK	Task	8	White/Hispanic	Hispanic
R013106	RC	Literary Experience	8	Male/Female	Female
R013302	RI	Task	8	White/Black	White
R013305	RI	Task	8	Male/Female	Female
R013404	RJ	Task	12	White/Hispanic	Hispanic
R013406	RJ	Task	8 12	Male/Female Male/Female	Female Female
R013407	RJ	Task	8	White/Hispanic	White
R013608	RE	Literary Experience	12	Male/Female	Male
R013907	RK	Task	12	White/Black	Black
R014706	RM	Information	8	Male/Female	Female
R015503	RM	Information	12	Male/Female	Female
R015608	RN	Information	12	Male/Female	Male

---

* For each grade for which an item was administered, three comparisons were performed: Male/Female, White/Black, and White/Hispanic.

Table L-2
Mathematics Items Identified as "C" Items in at least One Comparison*

Grade	Block	Scale	Grade	Comparison	Group Favored
M012131	MH	Algebra and Functions	12	White/Black	Black
M013031	MH	Data Analysis, Probability & Statistics	8	White/Black	White
M013331	MH	Measurement	8	White/Black	White
M013631	MH	Numbers and Operations	8	Male/Female	Male
M017401	MD	Numbers and Operations	12	Male/Female	Female
M017601	MD	Geometry	4	White/Black	Black
M020301	MF	Measurement	8	White/Black	White
			12	White/Black	White
M020501	MF	Numbers and Operations	12	Male/Female	Male
M021502	MF	Data Analysis, Probability & Statistics	12	White/Black	Black
			12	White/Hispanic	Hispanic
M022201	ME	Geometry	4	White/Black	White
M022701	ME	Numbers and Operations	8	Male/Female	Female
M022802	ME	Measurement	4	White/Black	White
M023601	ME	Data Analysis, Probability & Statistics	12	White/Black	White
M031701	MP	Estimation	4	White/Black	White
M032001	MP	Estimation	4	White/Black	White
			8	White/Black	White
M032801	MP	Estimation	4	White/Black	Black
M039501	MC	Measurement	4	White/Hispanic	White
M041501	ML	Numbers and Operations	4	Male/Female	Male
M042901	MM	Numbers and Operations	4	White/Black	Black
M043501	MM	Algebra and Functions	4	Male/Female	Female
M047101	Mk	Measurement	8	White/Black	White
M047901	MK	Measurement	8	Male/Female	Male
M048501	MO	Measurement	8	Male/Female	Male

---

* For each grade for which an item was administered, three comparisons were performed: Male/Female, White/Black, and White/Hispanic.

Table L-2 (continued)
Mathematics Items Identified as "C" Items in at least One Comparison*

Grade	Block	Scale	Grade	Comparison	Group Favored
M049101	MO	Numbers and Operations	8	Male/Female	Male
M050201	MC	Data Analysis, Probability & Statistics	8	White/Black	Black
M052401	MI	Numbers and Operations	12	Male/Female	Female
M052501	MI	Algebra and Functions	8	White/Black	Black
M052801	MI	Data Analysis, Probability & Statistics	8	White/Black	Black
M053501	ML	Algebra and Functions	8	White/Black	White
M054801	MN	Numbers and Operations	12	Male/Female	Female
M055201	MN	Numbers and Operations	8	White/Black	White
M055301	MN	Geometry	8	White/Black	Black
M055501	MN	Numbers and Operations	8	Male/Female	Female
M061904	MJ	Geometry	4	Male/Female	Male
M061907	MJ	Measurement	12	White/Black	White
M062501	MQ	Complex Problem Solving Skills	4	White/Black	White
			8	White/Black	White
			12	White/Black	White
M062701	MQ	Complex Problem Solving Skills	4	White/Black	White
			8	White/Black	White
M063101	MQ	Complex Problem Solving Skills	8	White/Hispanic	White
M063201	MQ	Complex Problem Solving Skills	12	Male/Female	Female
M063403	MQ	Complex Problem Solving Skills	8	White/Hispanic	White
M063404	MQ	Complex Problem Solving Skills	8	White/Hispanic	White
M064301	MR	Complex Problem Solving Skills	4	White/Black	Black
M064501	MR	Complex Problem Solving Skills	4	Male/Female	Male
			4	White/Black	White
M065301	MR	Complex Problem Solving Skills	8	White/Hispanic	Hispanic
N250231	MH	Data Analysis, Probability & Statistics	4	White/Black	White

---

* For each grade for which an item was administered, three comparisons were performed: Male/Female, White/Black, and White/Hispanic.

APPENDIX M

REANALYSIS OF THE 1990 MATHEMATICS ASSESSMENT DATA

Appendix M

Reanalysis of the 1990 Mathematics Assessment Data

John Mazzeo
Educational Testing Service

As has been the case since 1984, the estimation of proficiency scale distributions for the 1992 national and Trial State Assessments was carried out using the plausible values methodology described in Mislevy (1991) and in Chapter 8 of the current report. The methodology is implemented using Sheehan's (1985) MGROUP computer program and involves the estimation of a multivariate linear model for the regression of proficiency (θ) on a large number of predictor variables related to examinee background characteristics and instructional experience. The version of the program used in 1990, based on the EM algorithm described in Mislevy (1985), used Monte Carlo integration procedures to estimate the parameters of linear regression model. Subsequent to the 1990 assessment, these estimation procedures were improved by the introduction of analytic integration procedures and the incorporation of higher-order asymptotic corrections to estimates of examinee means and posterior variances (Thomas, 1992).

Preliminary research with simulated data and experience with selected reanalyses of previously reported 1990 NAEP data sets (both national and Trial State assessments) suggested that results from the revised program would differ from those obtained with earlier versions of MGROUP to a degree that was not ignorable. The 1990 estimates of the correlations between scales had been substantially attenuated. For example, estimates of correlations among the mathematics subscales (conditional on the full set of background variables used for the analysis) obtained with the version of MGROUP used in 1990 were typically in the .15 to .25 range across all states that participated in that year's Trial State Assessment. When the 1990 data were reanalyzed with the revised MGROUP procedures used for the 1992 Trial State Assessment, these same correlations were typically found to be in the .85 to .95 range.

The underestimation of subscale correlations had little impact on the accuracy of results (means, standard deviations, and percentiles) reported for the five NAEP content area scales. In addition, this underestimation had little effect on the means reported for the mathematics composite scale. However, the composite scale is a weighted average of the results from each of its constituent scales and, as such, the standard deviation of the composite is partly a function of the interscale correlations. Consequently, composite scale variability had been underestimated for both the national and Trial State Assessments and this attenuation resulted in underestimates of the percentages of examinees outside the more extreme NAEP anchor points.

Plans for the 1992 mathematics assessment called for the use of the revised estimation procedures. However, the use of such procedures for the 1992 analyses alone would make accurate comparisons to 1990 difficult, if not impossible, for certain composite scale statistics. In order to maintain the integrity of the 1990 NAEP mathematics scales for trend analysis, a decision was made to reanalyze the 1990 results for both the national and Trial State assessments and to report the revised 1990 figures in conjunction with the 1992 results. The reanalyses involved only those aspects related to the MGROUP procedure. The item parameter estimates from the 1990 national and Trial State assessments were *not* re-estimated. However, the estimation of conditioning models and the generation of plausible values were redone using the 1990 data and applying the same version of the MGROUP program used for the 1992 assessment. In all other respects, the 1990 reanalyses involved procedures nearly identical[1] to those that produced the original 1990 results (Yamamoto & Jenkins, 1992, section 13.2.6; Mazzeo, 1992, section 10.5).

Resetting the Origin and Unit of the 1990 Reporting Scales

The reanalysis of the 1990 results engendered slightly different scaling transformation constants—that is, the constants that linearly transform NAEP results from the metric in which they are estimated (θ_{est}, with approximate mean 0 and standard deviation of 1) to the proficiency metric in which they are reported (θ_{rpt}, with, in most cases, mean of 250.5 and standard deviation of 50. The procedures used to obtain scaling transformation constants for the revised 1990 results are described below.

Scaling Transformations for the 1990 National Mathematics Assessment

The revised 1990 scaling transformations were obtained by applying the same procedures originally used in 1990 (Yamamoto & Jenkins, 1992) to the revised 1990 results.

For the Numbers and Operations, Measurement, Geometry, and Algebra and Functions scales the procedure was as follows:

1) Separate estimates of the mean and variance of each scale were obtained in the national θ-metric for the winter half-sample of each of the three age/grade cohorts. These estimates were obtained using final sampling weights. The reason for using only the winter half-sample was to center the scale and establish its unit in terms of the most sensible reference population against which to compare the Trial State Assessment. Participants in the Trial State Assessment were tested during the same time period as the winter half-sample of the 1990 national assessment (January through March of 1990). Note, however, that this national standardization population is still not directly comparable to the Trial State Assessment population in that a) it contains age-eligible as well as grade-eligible examinees, b) contains private-school as well as public-school examinees, and c) contains examinees from states that did not participate in Trial State Assessment. It should also be noted that the national sample to which the Trial State Assessment is compared in the 1990 composite and state reports is not identical

---

[1] A slight change to the program used to obtain principal components resulted in one additional component being included as a predictor variable in the conditioning model for each age/grade cohort of the national assessment and for each jurisdiction of the Trial State Assessment.

to the national standardization sample. The national reporting comparison sample excluded age-eligible and private-school examinees.

2) The estimates from step 1 were combined to produce an estimate of the overall mean and standard deviation for a population of students tested in the winter and consisting of equal numbers of students from each age/grade cohort. The estimate of the overall mean was simply the unweighted average of the means for the three cohorts. The estimate of the standard deviation was obtained as the square root of the sum of the unweighted average of the within-cohort variances and the variance of the between-cohort means.

3) Constants were then derived that linearly transform the overall mean and standard deviation obtained in step 2 to 250.5 and 50, respectively. These values, which are consistent with previous NAEP practice, represent the mean and standard deviation of estimated true scores for the combined population of all three cohorts on a hypothetical test. The test consists of 500 items, equally-spaced between -4.99 and 4.99 on the θ_{est} scale, which follow a 1-parameter logistic model with a discrimination parameter of 1.5 (see Beaton, 1987, page 384).

Sufficient items to produce a Data Analysis, Statistics, and Probability scale were present for only two of the three age/grade cohorts. Therefore, a slightly different scale transformation procedure was used for this scale in 1990 and with the revised 1990 results. The procedure was as follows:

1) For the Data Analysis scale, means were obtained in the national θ metric for the winter half-samples of the grade 8/age 13 and grade 12/age 17 cohorts.

2) For the other four scales discussed above, means were obtained in the reporting metric for these same cohort.

3) For each cohort, a weighted average of the means obtained in step 2 was produced. The weights used were those employed in forming the 1990 composite scale at the corresponding grade, renormalized to sum to 1.

4) Constants were then derived that mapped the θ metric means obtained in step 1 to the weighted averages produced in step 3.

The procedure used to reset the Estimation scale was identical to that used to set the metric for the Numbers and Operations, Measurement, Geometry, and Algebra and Functions scales, the only difference being the use of both the winter and spring age/grade samples. The decision to use both half-samples was based on two considerations. First, because the estimation items had not been administered in the 1990 Trial State Assessment there was less need to center the scales in terms of a comparable reference sample. Second, the estimation items were administered to a separate, and smaller, national sample than was the BIB portion of the main assessment (which contains the items for the other scales). Restricting data to the winter-half sample for the estimation scale would have resulted in smaller than desired sample sizes for the standardization.

The original 1990 transformations and their corresponding revised values are given in Table M-1. The transformations are of the form $\theta_{rpt} = k_1 + k_2(\theta_{cal})$.

Scaling Transformations for the 1990 Trial State Assessment in Mathematics

The scaling transformations for the 1990 Trial State Assessment were also redone using identical procedures to those reported in Mazzeo (1991), but applied to the reestimated 1990 results. The procedure, which was intended to equate the metrics of the Trial State Assessment scales to their corresponding national scales, was as follows:

1) Means and standard deviations in the Trial State Assessment θ-metric were obtained for each of the five scales for the aggregate sample of Trial State Assessment examinees from all participating jurisdictions with the exception of Guam and the Virgin Islands. These latter two participants were excluded because no data from corresponding PSUs were available from the national assessment. Final sampling weights provided by Westat were used in producing the necessary sample moments.

2) Corresponding means and standard deviations for these scales in the national θ-metric were also obtained for a restricted sample of the national assessment, referred to as the State Aggregate Comparison (SAC) sample. The SAC sample consisted of public-school, grade-eligible students from only those states that participated in the 1990 Trial State Assessment. Special weights were provided by Westat to enhance comparability of this sample to the aggregate of the Trial State Assessment on which the means and standard deviations from step 1 were obtained.

3) A set of transformation constants were obtained linking the two θ-metrics by setting means and standard deviations equal.

4) These constants were then concatenated with the appropriate constants in Table 1 to produce a final set of revised Trial State Assessment scaling constants.

The original and revised Trial State Assessment scaling constants are given in Table M-2.

Comparisons of Original and Revised Results for the 1990 Trial State Assessment

In the vast majority of cases, differences between the 1990 results originally reported and the revised results are extremely small. For example, the revised state means and means for subgroups on each of the five content area scales and the composite scales are, for the most part, within 1 standard error of the originally reported values. Figure M-1 provides plots of the revised state means versus original means for the Numbers and Operations scale (the scale for which examinees were administered the greatest number of items), the Data Analysis, Statistics, and Probability scale (the scale for which examinees were administered the fewest items), and

Table M-1
Original and Revised Scaling Transformation Constants for the 1990 National Mathematics Scales

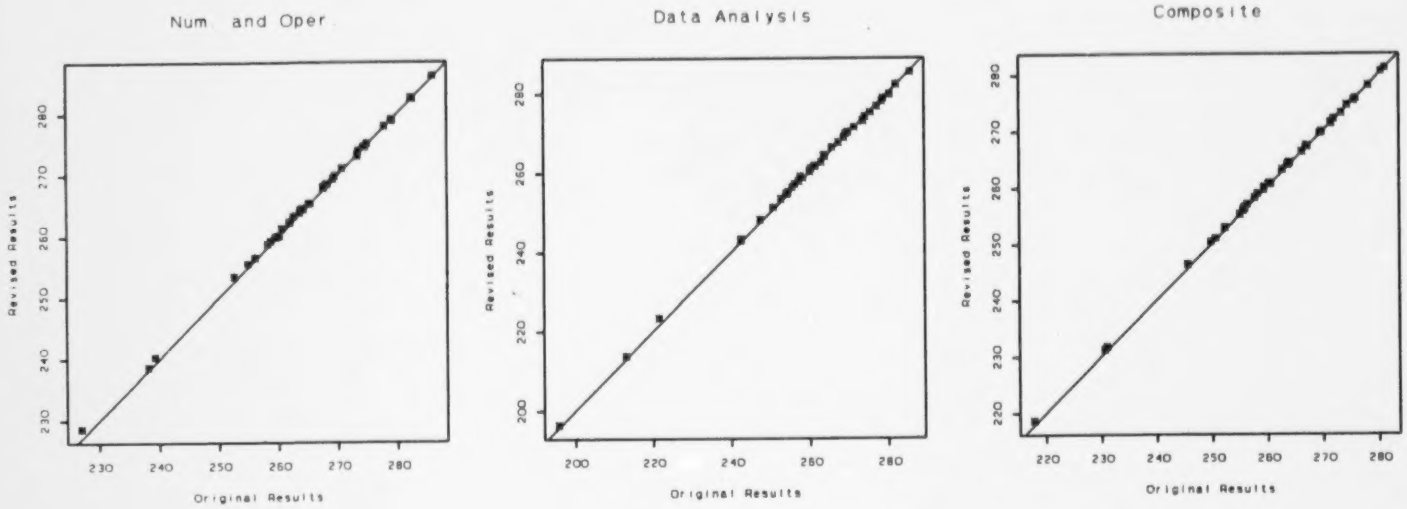
	Original Transformation		Revised Transformation	
Scale	**k_1**	**k_2**	**k_1**	**k_2**
Numbers & Operations	251.72	50.35	251.82	50.12
Measurement	252.59	49.99	252.18	49.79
Geometry	252.58	49.15	252.79	48.81
Data Analysis	274.08	44.47	274.19	43.68
Algebra & Functions	252.01	50.34	251.90	49.99
Estimation	249.52	49.62	249.51	49.53

Table M-2
Original and Revised Scaling Transformation Constants for the 1990 Trial State Assessment Mathematics Scales

	Original Transformation		Revised Transformation	
Scale	**k_1**	**k_2**	**k_1**	**k_2**
Numbers & Operations	265.28	36.14	265.87	35.35
Measurement	256.69	43.92	256.77	42.18
Geometry	258.90	35.81	259.23	34.79
Data Analysis	259.64	44.84	260.21	43.09
Algebra & Functions	259.71	38.12	260.27	37.43

Figure M-1

Revised 1990 State Means Plotted Against Originally Reported 1990 State Means for Three Scales



for the composite scale for each of the 40 participants. As is evident from the figure, differences were quite small and the rank ordering of the states was unaffected. Essentially identical results were obtained for the scales not shown.

As discussed earlier, the principal difference between the results produced by the MGROUP program used in 1990 and that used in 1992 involved estimates of the within-jurisdiction standard deviations and the proportions of students exceeding NAEP anchor points. Figure M-2 contains a stem-and-leaf display of the ratio of the revised 1990 composite scale standard deviation estimates to the original estimates for each of the 40 participants. The results originally reported were consistently lower than the revised results, with the underestimation ranging from 6 to 16 percent.

Figure M-2
Stem-and-Leaf Display* of
Ratios of Composite Standard Deviations (Revised 1990/Original 1990)

N = 40, Median = 1.1075, Quartiles = 1.0945, 1.121
Decimal point is 2 places to the left of the colon

```
 1    1  106 : 3
 2    1  107 : 1
 8    6  108 : 224558
13    5  109 : 13689
     10  110 : 2356777899
17    5  111 : 01589
12    8  112 : 00235578
 4    2  113 : 02
 2    1  114 : 1
 1    0  115 :
 1    1  116 : 3
```

---

* The first column of numbers shows observation depths; the second column shows the number of observations; the remainder of the figure contains the stem-and-leaf display.

Table M-3 provides differences (original 1990 results minus revised 1990 results) in the estimated proportions at or above each NAEP anchor point for each of the participating jurisdictions. Differences were typically on the order of 1 to 2 percent and none exceeded 4 percent. In general, the percentages at or above the higher anchor points were slightly underestimated, while the percentages at or above the lowest anchor point were slightly overestimated. These differences were a direct result of the underestimation of the correlations between scales and the resulting underestimation in the composite scale standard deviations.

Table M-3
Differences (Original 1990 Results Minus Revised 1990 Results) in Estimated Percentages At or Above Each NAEP Anchor Point for Each Participating Jurisdiction

Jurisdiction	**NAEP Anchor Levels**			
	200	**250**	**300**	**350**
Alabama	1.8	-1.1	-1.8	-0.2
Arizona	2.1	-0.3	-2.2	-0.3
Arkansas	1.8	-0.6	-2.0	-0.1
California	2.4	-1.0	-1.4	-0.3
Colorado	1.6	1.5	-2.2	-0.3
Connecticut	1.3	1.0	-2.3	-0.4
Delaware	1.9	-0.7	-0.8	-0.2
District of Columbia	3.1	-3.1	-0.6	-0.1
Florida	2.5	-1.5	-1.3	-0.2
Georgia	1.5	-0.9	-1.5	-0.4
Guam	2.6	-3.8	-0.7	-0.1
Hawaii	2.8	-2.1	-1.1	-0.3
Idaho	0.7	2.7	-2.6	-0.2
Illinois	1.7	0.9	-2.5	-0.4
Indiana	1.0	1.2	-2.3	-0.4
Iowa	0.5	3.2	-3.1	-0.4
Kentucky	1.9	-0.6	-2.1	-0.2
Louisiana	2.2	-2.6	-1.3	-0.2
Maryland	1.7	-0.6	-2.0	-0.4
Michigan	1.1	0.3	-2.2	-0.6
Minnesota	0.7	2.7	-2.1	-0.8
Montana	0.3	2.6	-2.4	-0.6
Nebraska	1.0	2.0	-2.6	-0.5
New Hampshire	1.0	1.8	-2.3	-0.3
New Jersey	1.1	0.9	-1.7	-0.5
New Mexico	1.9	-1.1	-1.9	-0.2
New York	2.0	-0.4	-1.9	-0.6
North Carolina	2.4	-1.9	-1.4	-0.1
North Dakota	0.3	2.4	-2.2	-0.8
Ohio	1.2	0.4	-2.0	-0.3
Oklahoma	1.4	0.4	-3.0	-0.1
Oregon	1.5	1.6	-1.6	-0.5
Pennsylvania	1.2	-0.2	-1.8	-0.4
Rhode Island	0.9	-0.5	-1.9	-0.2
Texas	1.8	-0.5	-2.0	-0.4
Virginia	1.1	-0.4	-1.5	-0.6
Virgin Islands	1.9	-3.0	-0.2	-0.0
West Virginia	2.2	-0.9	-1.4	-0.2
Wisconsin	0.6	2.2	-2.7	-0.3
Wyoming	0.8	2.2	-3.1	-0.1

APPENDIX N

Estimation Error Variance by Gender and Race/ethnicity

Table N-1
Estimation Error Variance
for the Main Reading for Literary Experience Scale

		Proportion of Variance Due to	
	Total Variance	**Student Sampling**	**Latency of θ**
GRADE 4			
Total	0.90	0.86	0.14
Male	1.56	0.86	0.14
Female	1.12	0.76	0.24
White	1.41	0.87	0.13
Black	2.94	0.75	0.25
Hispanic	5.65	0.87	0.13
GRADE 8			
Total	0.73	0.79	0.21
Male	1.07	0.72	0.28
Female	1.05	0.77	0.23
White	1.09	0.81	0.19
Black	2.32	0.74	0.26
Hispanic	2.55	0.69	0.31
GRADE 12			
Total	0.59	0.58	0.42
Male	0.98	0.58	0.42
Female	0.94	0.60	0.40
White	0.73	0.49	0.51
Black	3.65	0.74	0.26
Hispanic	11.23	0.88	0.12

Table N-2
Estimation Error Variance
for the Main Reading to Gain Information Scale

		Proportion of Variance Due to	
	Total Variance	**Student Sampling**	**Latency of θ**
GRADE 4			
Total	1.19	0.89	0.11
Male	1.81	0.89	0.11
Female	1.38	0.79	0.21
White	1.87	0.90	0.10
Black	3.20	0.83	0.17
Hispanic	3.88	0.81	0.19
GRADE 8			
Total	0.63	0.85	0.15
Male	0.88	0.82	0.18
Female	0.90	0.83	0.17
White	0.92	0.86	0.14
Black	2.26	0.82	0.18
Hispanic	1.99	0.79	0.21
GRADE 12			
Total	0.35	0.86	0.14
Male	0.53	0.76	0.24
Female	0.57	0.84	0.16
White	0.47	0.85	0.15
Black	2.16	0.87	0.13
Hispanic	4.05	0.88	0.12

Table N-3
Estimation Error Variance
for the Main Reading to Perform a Task Scale

	Total Variance	Student Sampling	Latency of θ
	——— Proportion of Variance Due to ———		
GRADE 8			
Total	0.72	0.76	0.24
Male	1.00	0.72	0.28
Female	1.02	0.74	0.26
White	1.01	0.78	0.22
Black	2.47	0.67	0.33
Hispanic	2.87	0.72	0.28
GRADE 12			
Total	0.43	0.70	0.30
Male	0.76	0.62	0.38
Female	0.69	0.66	0.34
White	0.54	0.64	0.36
Black	2.56	0.75	0.25
Hispanic	5.39	0.79	0.21

Table N-4
Estimation Error Variance
for the Main Reading Composite Scale

	Total Variance	Student Sampling	Latency of θ
		Proportion of Variance Due to	
GRADE 4			
Total	0.89	0.94	0.06
Male	1.48	0.93	0.07
Female	1.01	0.86	0.14
White	1.42	0.95	0.05
Black	2.55	0.86	0.14
Hispanic	3.98	0.89	0.11
GRADE 8			
Total	0.56	0.92	0.08
Male	0.77	0.90	0.10
Female	0.79	0.91	0.09
White	0.85	0.92	0.08
Black	1.90	0.84	0.16
Hispanic	1.86	0.86	0.14
GRADE 12			
Total	0.35	0.84	0.16
Male	0.54	0.78	0.22
Female	0.54	0.84	0.16
White	0.43	0.79	0.21
Black	2.25	0.89	0.11
Hispanic	5.64	0.93	0.07

Table N-5
Estimation Error Variance
for the Reading Long-term Trend Scale

		Proportion of Variance Due to	
	Total Variance	Student Sampling	Latency of θ
AGE 9			
Total	0.73	0.88	0.12
Male	1.37	0.88	0.12
Female	0.84	0.83	0.17
White	0.89	0.90	0.10
Black	4.33	0.82	0.18
Hispanic	9.64	0.93	0.07
AGE 13			
Total	1.30	0.91	0.09
Male	2.51	0.90	0.10
Female	1.49	0.85	0.15
White	1.48	0.92	0.08
Black	5.53	0.84	0.16
Hispanic	11.45	0.87	0.13
AGE 17			
Total	1.20	0.86	0.14
Male	2.21	0.85	0.15
Female	1.45	0.77	0.23
White	1.77	0.88	0.12
Black	4.01	0.72	0.28
Hispanic	13.72	0.85	0.15

Table N-6
Estimation Error Variance
for the Main Numbers and Operations Scale

	Total Variance	Proportion of Variance Due to: Student Sampling	Proportion of Variance Due to: Latency of θ
GRADE 4			
Total	0.68	0.87	0.13
Male	0.88	0.80	0.20
Female	1.30	0.89	0.11
White	0.91	0.88	0.12
Black	2.25	0.81	0.19
Hispanic	2.42	0.83	0.17
GRADE 8			
Total	0.71	0.91	0.09
Male	1.14	0.91	0.09
Female	0.92	0.86	0.14
White	0.86	0.90	0.10
Black	2.07	0.81	0.19
Hispanic	1.78	0.75	0.25
GRADE 12			
Total	0.70	0.85	0.15
Male	1.07	0.83	0.17
Female	0.96	0.82	0.18
White	0.73	0.83	0.17
Black	2.64	0.85	0.15
Hispanic	2.66	0.75	0.25

Table N-7
Estimation Error Variance
for the Main Measurement Scale

		Proportion of Variance Due to	
	Total Variance	Student Sampling	Latency of θ
GRADE 4			
Total	0.75	0.83	0.17
Male	1.02	0.78	0.22
Female	1.26	0.79	0.21
White	1.05	0.84	0.16
Black	2.75	0.74	0.26
Hispanic	2.22	0.72	0.28
GRADE 8			
Total	1.38	0.85	0.15
Male	1.96	0.86	0.14
Female	1.93	0.80	0.20
White	1.69	0.84	0.16
Black	4.11	0.75	0.25
Hispanic	3.02	0.69	0.31
GRADE 12			
Total	0.93	0.84	0.16
Male	1.50	0.85	0.15
Female	1.29	0.79	0.21
White	1.05	0.82	0.18
Black	3.93	0.82	0.18
Hispanic	3.47	0.69	0.31

Table N-8
Estimation Error Variance
for the Main Geometry Scale

		Proportion of Variance Due to	
	Total Variance	**Student Sampling**	**Latency of θ**
GRADE 4			
Total	0.49	0.78	0.22
Male	0.76	0.73	0.27
Female	0.90	0.75	0.25
White	0.76	0.78	0.22
Black	2.39	0.69	0.31
Hispanic	1.98	0.68	0.32
GRADE 8			
Total	0.88	0.86	0.14
Male	1.28	0.85	0.15
Female	1.09	0.82	0.18
White	1.30	0.87	0.13
Black	2.21	0.72	0.28
Hispanic	1.69	0.67	0.33
GRADE 12			
Total	1.01	0.86	0.14
Male	1.67	0.83	0.17
Female	1.33	0.83	0.17
White	1.13	0.82	0.18
Black	4.01	0.86	0.14
Hispanic	5.29	0.83	0.17

Table N-9
Estimation Error Variance
for the Main Data Analysis, Statistics, and Probability Scale

	Total Variance	Proportion of Variance Due to — Student Sampling	Latency of θ
GRADE 4			
Total	0.67	0.81	0.19
Male	0.98	0.72	0.28
Female	1.23	0.78	0.22
White	0.95	0.81	0.19
Black	2.42	0.71	0.29
Hispanic	2.32	0.70	0.30
GRADE 8			
Total	1.15	0.85	0.15
Male	1.79	0.85	0.15
Female	1.56	0.79	0.21
White	1.38	0.84	0.16
Black	3.01	0.73	0.27
Hispanic	2.20	0.68	0.32
GRADE 12			
Total	1.05	0.81	0.19
Male	1.49	0.81	0.19
Female	1.32	0.76	0.24
White	1.17	0.79	0.21
Black	3.73	0.80	0.20
Hispanic	4.18	0.67	0.33

Table N-10
Estimation Error Variance
for the Main Algebra and Functions Scale

	Proportion of Variance Due to		
	Total Variance	**Student Sampling**	**Latency of θ**
GRADE 4			
Total	0.75	0.76	0.24
Male	1.11	0.63	0.37
Female	1.40	0.74	0.26
White	1.12	0.75	0.25
Black	2.87	0.64	0.36
Hispanic	2.56	0.63	0.37
GRADE 8			
Total	1.01	0.85	0.15
Male	1.52	0.84	0.16
Female	1.30	0.80	0.20
White	1.38	0.84	0.16
Black	2.24	0.71	0.29
Hispanic	2.27	0.70	0.30
GRADE 12			
Total	0.99	0.94	0.06
Male	1.42	0.91	0.09
Female	1.30	0.91	0.09
White	1.04	0.91	0.09
Black	4.49	0.92	0.08
Hispanic	3.07	0.81	0.19

Table N-11
Estimation Error Variance
for the Main Mathematics Composite Scale

		Proportion of Variance Due to	
	Total Variance	Student Sampling	Latency of θ
GRADE 4			
Total	0.55	0.95	0.05
Male	0.70	0.91	0.09
Female	1.00	0.95	0.05
White	0.78	0.95	0.05
Black	1.81	0.88	0.12
Hispanic	1.74	0.89	0.11
GRADE 8			
Total	0.80	0.98	0.02
Male	1.23	0.96	0.04
Female	1.00	0.96	0.04
White	1.03	0.97	0.03
Black	1.86	0.90	0.10
Hispanic	1.44	0.89	0.11
GRADE 12			
Total	0.77	0.97	0.03
Male	1.17	0.96	0.04
Female	1.00	0.96	0.04
White	0.81	0.97	0.03
Black	3.14	0.95	0.05
Hispanic	2.80	0.91	0.09

Table N-12
Estimation Error Variance for the Main Estimation Scale

	Proportion of Variance Due to		
	Total Variance	Student Sampling	Latency of θ
GRADE 4			
Total	2.44	0.79	0.21
Male	2.75	0.70	0.30
Female	4.58	0.76	0.24
White	3.47	0.80	0.20
Black	10.43	0.70	0.30
Hispanic	9.01	0.64	0.36
GRADE 8			
Total	1.58	0.89	0.11
Male	2.03	0.83	0.17
Female	2.19	0.83	0.17
White	1.90	0.87	0.13
Black	10.31	0.87	0.13
Hispanic	6.52	0.82	0.18
GRADE 12			
Total	1.27	0.89	0.11
Male	1.68	0.83	0.17
Female	2.10	0.86	0.14
White	1.67	0.86	0.14
Black	2.88	0.73	0.27
Hispanic	7.64	0.81	0.19

Table N-13
Estimation Error Variance
for the Mathematics Long-term Trend Scale

	Total Variance	Proportion of Variance Due to	
		Student Sampling	Latency of θ
AGE 9			
Total	0.70	0.91	0.09
Male	0.90	0.83	0.17
Female	0.91	0.87	0.13
White	0.67	0.88	0.12
Black	4.02	0.89	0.11
Hispanic	5.10	0.88	0.12
AGE 13			
Total	0.70	0.94	0.06
Male	0.81	0.89	0.11
Female	1.02	0.92	0.08
White	0.87	0.94	0.06
Black	3.34	0.91	0.09
Hispanic	3.35	0.81	0.19
AGE 17			
Total	0.79	0.93	0.07
Male	1.21	0.90	0.10
Female	1.20	0.91	0.09
White	0.73	0.90	0.10
Black	4.98	0.88	0.12
Hispanic	7.49	0.85	0.15

Table N-14
Estimation Error Variance
for the Science Long-term Trend Scale

	Proportion of Variance Due to		
	Total Variance	**Student Sampling**	**Latency of θ**
AGE 9			
Total	0.70	0.81	0.19
Male	1.19	0.78	0.22
Female	1.08	0.75	0.25
White	0.80	0.84	0.16
Black	4.72	0.64	0.36
Hispanic	6.32	0.83	0.17
AGE 13			
Total	0.71	0.87	0.13
Male	1.23	0.82	0.18
Female	1.06	0.83	0.17
White	0.80	0.85	0.15
Black	4.55	0.79	0.21
Hispanic	6.24	0.81	0.19
AGE 17			
Total	1.66	0.92	0.08
Male	2.89	0.89	0.11
Female	2.03	0.88	0.12
White	1.62	0.90	0.10
Black	10.47	0.81	0.19
Hispanic	30.77	0.94	0.06

Table N-15
Estimation Error Variance for the Main Writing Scale

		Proportion of Variance Due to	
	Total Variance	**Student Sampling**	**Latency of θ**
GRADE 4			
Total	1.01	0.68	0.32
Male	2.00	0.69	0.31
Female	1.32	0.62	0.38
White	1.46	0.68	0.32
Black	3.94	0.66	0.34
Hispanic	4.17	0.63	0.37
GRADE 8			
Total	0.64	0.65	0.35
Male	1.14	0.55	0.45
Female	0.85	0.63	0.37
White	0.83	0.66	0.34
Black	3.19	0.70	0.30
Hispanic	2.86	0.54	0.46
GRADE 12			
Total	0.94	0.76	0.24
Male	1.20	0.70	0.30
Female	1.50	0.72	0.28
White	1.25	0.75	0.25
Black	4.17	0.75	0.25
Hispanic	6.26	0.72	0.28

Table N-16
Estimation Error Variance
for the Writing Long-term Trend Scale

		Proportion of Variance Due to	
	Total Variance	Student Sampling	Latency of θ
GRADE 4			
Total	1.40	0.43	0.57
Male	2.61	0.42	0.58
Female	2.32	0.43	0.57
White	1.68	0.49	0.51
Black	10.93	0.44	0.56
Hispanic	9.96	0.43	0.57
GRADE 8			
Total	1.81	0.61	0.39
Male	2.78	0.46	0.54
Female	2.75	0.62	0.38
White	1.96	0.53	0.47
Black	12.51	0.61	0.39
Hispanic	13.62	0.46	0.54
GRADE 11			
Total	1.11	0.45	0.55
Male	1.92	0.44	0.56
Female	1.78	0.35	0.65
White	1.26	0.48	0.52
Black	6.82	0.27	0.73
Hispanic	12.76	0.58	0.42

APPENDIX O

CORRESPONDENCE RELATED TO THE ANALYSIS OF THE 1992 MATHEMATICS SHORT-TERM TREND DATA

BOSTON COLLEGE
CHESTNUT HILL, MASSACHUSETTS 02167

Prof. Albert E. Beaton
Center for the Study of Testing,
Evaluation, and Educational Policy
320 Campion Hall
Phone (617)552-4638
FAX (617)552-8419

October 29, 1992

Dr. Eugene Johnson
National Assessment of Educational Progress
T-02
Educational Testing Service
Princeton, NJ 08541

Dear Gene:

Thank you for the opportunity to review the 1992 NAEP results. I suggest that more research be done on their meaning before the results are published.

As you know, the problem is as follows: there has been a substantial increase in NAEP performance in mathematics at all grade levels, and this increase is large enough to suggest that some other factor or factors may have influenced the student performance. I realize that interpolation and extrapolation on the NAEP scale is not fully justified, but it does give us a reasonable yardstick by which to measure the size of the gains. If my calculations are correct, the fourth graders went up 43% of a year's growth, the eighth graders went up 64%, and, by extrapolation, the twelfth graders went up 56% of a year's growth. These gains are not as large as the twelfth grade drop in the 1986 reading anomaly, but are still large enough to merit further investigation. I understand that performance in the other subject areas has also risen.

On my trip to ETS, I reviewed the work that you have done until now. This work has used only the 1990 and 1992 data short-term trend data and thus does not necessarily reflect long-term trends. Of particular interest was the partitioning analyses that showed the rise to be nearly universal, occurring in all sections of the country and for gender, racial/ethnic, and other groupings. There was one substantial exception: the student non-response adjustment factor. Clearly, this should be investigated by your group or Westat to assure that no biasing errors occurred in creating this variable.

It is also important to look at these changes in relation to the variance of past results as evidenced by the long term trends. My memory suggests that these gains are much larger than any changes reported in the past.

Since the phenomenon of rising scores is reflected in the proportion-correct statistics as well as the scale values, we do not need to review the scaling technology as a potential factor.

I am afraid that I have come up with an hypothesis as to what has happened, that is, the higher stakes and visibility of NAEP are doing what theory dictates: raising the scores. I sincerely hope, for NAEP's sake, that we can reject this hypothesis or find a way too isolate and adjust for the phenomenon.

There are several observations that suggest this hypothesis:

1) The percentage of non-cooperating schools has increased slightly. Consider the possibility that the lower-scoring schools are just plain more unwilling to cooperate and more of them refuse. I would expect this to happen more in the TSA schools that the national schools, but I doubt that some of the school people would distinguish between them carefully. In this case, the average scores would rise slightly because of fewer low-scoring schools.

2) The percentage of excluded students has risen slightly. Again, by excluding a few extra low-scoring students, the average would rise, even though the population as a whole had not changed.

3) The students try harder and thus score better. In this case, it would be the previous year(s) that underestimated what students knew and could do.

* * *

We can get an idea of the effect of points 1 and 2 by a straightforward calculation which involves making inferences about a full population in which there were no non-cooperating schools and no excluded students. Under a reasonable assumption, we can compute selected percentiles of the 1990 and 1992 full distributions and their trimmed means. The 1990 and 1992 score distributions can then be compared after this adjustment for changes in participation.

The basic idea is as follows:

Let us assume that the students from the non-cooperating schools and the excluded students would have scored below the 25th percentile, if they had taken the test. This is not an unreasonable assumption, nor is it very precise since it is not a point estimate: the students could have fallen anywhere below the 25th percentile. Let k be the percentage of students are missing or excluded, defined as a percentage of the students who were assessed. Then,

N_a is the assessed population. (I am using population just to indicate that sampling weights should be used. See below for a comment on sampling weights.) For simplicity, we can set this to 100.

N_b is the estimated number of students in the population who would have been assessed if participation had been complete and no child was excluded. That is, $N_b=.01*(100+k)*N_a$.

Under the assumption that the k missing students would score below the 25th percentile *of those who were assessed*, the 25th percentile of the score distribution of the N_b students would be at

$Q_b(.25*N_b)=Q_a(.25*N_b-k)$

where Q_x is the score of the student at the position in the distribution x defined by the term in parentheses. Thus, the point below which 25% of the students in the full distribution fall is the same as the point below which a smaller percentage of students in the assessed distribution fall.

Let us say that 8% of the students were not assessed so that N_b is 108. then,

$Q_b(.25*108)=Q_a(.25*108-k)$

$=Q_b(26)=Q_a(18)$

That is, the 26th person in the ordered full distribution would be at its 25th %ile and this would be at the 18th %ile in the assessed population, since the missing 8 students are assumed to score lower.

This formula generalizes for any percentile by substituting the proportion (p) below that point for .25,

$Q_b(p*N_b)=Q_a(p*N_b-k)$

The assumption is that the missing or excluded students would fall somewhere below that percentile value.

For more simple examples, if $N_a=100$ and $N_b=125$, then k=25 and the

25th %ile in the full population ==> 6.25th %ile in the assessed,
50th %ile in the full population ==> 37.5th %ile in the assessed, and
75th %ile in the full population ==> 68.75th %ile in the assessed.

Using these values for the assessed distribution, you can compare two distributions, either by percentiles or by trimming those cases below the 25th and above the 75th percentiles. As all Tukey students know, there are many reasons for using the trimmed mean as a single statistic for comparing distributions.

In NAEP, this comparison can be applied directly to the excluded students but not to the students whose schools did not cooperate since the weights are already adjusted for non-response. This comparison would have to be applied to the unadjusted weights.

This simple procedure should allow you to estimate the gross effect of non-participation and increased exclusion. I suspect that this adjustment will remove some, but not all, of the increase in performance.

* * *

The evidence is unclear on the hypothesis that students tried harder and did better in 1992 than 1990. From some tables that I reviewed, the students reached more items and answered more correctly at grades four and eight but not at grade twelve. John Mazzeo showed some data that contradicted the tables. I think that this hypothesis is so important that the data should be carefully checked and reviewed again.

To make sense of this hypothesis, we would need to divide the 1990 students who did not reach an item into two groups, the lazy and the slow. The 1990 lazy students knew the answers but didn't try, although their 1992 counterparts did and answered the items correctly, raising the average scores. The slow students still wouldn't reach the end of the test, and so would not have an effect on the averages. This hypothesis is a little hard to believe, especially given the recent research demonstrating how hard it was to increase motivation for NAEP. But as long as the original charts show such changes in non-response, we cannot rule this hypothesis out.

* * *

I think that we have to be very careful, with NAEP, that we do not keep analyzing significant changes until we can throw them out for technical reasons. However, the changes in cooperation and exclusion rates will clearly affect reports and must be studied. Let us hope that adjusting for these changes removes just a little of the gains.

Sincerely,

Albert E. Beaton

cc: David Hobson
Archie Lapointe
Ina Mullis

EDUCATIONAL TESTING SERVICE PRINCETON, N.J. 08541

[illegible]

November 24, 1992

Professor Albert E. Beaton
Boston College
Center for the Study of Testing,
Evaluation, and, Educational Policy
320 Champion Hall
Chestnut Hill, Massachusetts 02167

Dear Al:

Thank you for your suggestions on lines of research about the 1992 NAEP mathematics results. After your trip to ETS we performed many of the analyses suggested in your October 29 letter. In particular, we have been digging into the effect of exclusion and nonresponse on the 1990 and 1992 results. Some results from our analyses to date on this are given in Tables 1 through 3.

Each of Tables 1 through 3 pertains to the national samples for a particular grade and compare 1992 results with results from the 1990 winter half sample. Each table is in three parts. Part A shows the percentiles computed using our usual methods (i.e. based only on nonexcluded students with nonresponse adjusted weights). Part A also shows the exclusion and absentee rates. Part B shows the percentiles recomputed by rolling in the excluded students according to the procedure given in your letter. Part C shows percentiles with adjustments (as in your letter) for both excluded and absent students. The base weights (no trimming, poststratification or nonresponse adjustments) were used in this case.

Comparing the between year differences from Parts A, B, and C suggests to us that the effects of exclusion and absenteeism on the change between 1990 and 1992 for grades 4 and 8 are modest, particularly for the 50th and higher percentiles and especially when the standard errors of the percentiles are considered. The effect of absenteeism is more dramatic at grade 12, but it is in the direction of increasing the change; the effect of exclusion is slight. We are currently working on assessing the effects of school nonresponse and will let you know the results.

Subsequent to your visit we spent some additional time checking and double-checking the Item Analysis results for the short-term trend blocks for the National and Trial State Assessments. As you undoubtedly remember, there was conflicting data from the National and Trial State Assessments regarding changes in the percents of students reaching the later items in these trend blocks.

After some additional investigation, we were able to determine that the National figures shown at the most recent DAC meeting were not quite correct. Tables 4 through 7 contain the corrected results. The tables show, for each block, the percentage of examinees not reaching the last item, the average percentage of not-reached items, and the average percent correct. Separate tables are given for the National Assessment at grades 4, 8, and 12, and for the Trial State Assessment at grade 8.

The National results compare the 1992 grade samples with the grade samples from the 1990 winter half sample. Our reading of the national results is that there was little difference in not-reached rates at grades 4 and 8, and a noticeable *increase* from 1990 to 1992 in not-reached rates at grade 12 for 3 of the 4 blocks. In other words, where differences exist, they are in the direction of 1992 examinees reaching fewer items.

The Trial State Assessment grade 8 results compare the aggregate sample from 1990 to the aggregate of the 1992 states that participated in both assessments. In both cases, the so-called Senate weights were used. There is a marked difference in not-reached rates for three of the four blocks. However, again the direction of the difference is that 1992 examinees reach fewer items.

We are not sure what to make of all this. One could argue that for the higher stakes Trial State Assessment, examinees are working more slowly and carefully. Therefore, they reach fewer items but do better on the items that they attempt. However, the grade 8 gains in percent-correct for the National Assessment (where there essentially no differences in not-reached rates) are about the same magnitude as those for the Trial State Assessment (where the change is substantial). Furthermore, block MF -- which showed almost no change from 1990 to 1992 in the percentage not-reached -- showed the largest increase in percent correct for the Trial State Assessment. In sum, these results suggest that changes in the Trial State Assessment not-reached rates probably account for little of the gains.

We did do some additional analyses comparing gains in item percent correct to changes in the percent of examinees reaching an item. Gains from 90 to 92 in percent-correct were evident even for items which appear early in a block (where almost everyone in both 1990 and 1992 attempted the item) and increase little, if at all, as differences in not-reached rates increased.

It appears that the data we have about not-reached rates does not suggest that changes in student motivation account for much, if any, of the gains. As you indicated in your letter of October 29, other research on the difficulty in increasing motivation in NAEP also seems to discount differential motivation as the culprit.

If you have any further thoughts or suggestions for analyses please let us know.

Sincerely,

Gene

Eugene G. Johnson

John Mazzeo

cc: Archie Lapointe
Ina Mullis
Frank Jenkins

Table 1: Percentiles of Overall Mathematics Proficiency under three treatments of excluded and absent students
Grade 4

	A: Percentiles for nonexcluded students with nonresponse adjustment for absent students*					Percent	
	25th	50th	75th	90th	95th	excluded	absent
1992	197.1(1.0)	220.1(1.0)	241.1(1.2)	258.9(0.8)	269.4(1.5)	6.7	6.2
1990	192.8(1.0)	213.9(1.1)	234.7(1.1)	253.2(1.5)	263.9(1.7)	5.2	6.8
difference	4.3	6.1	6.4	5.7	5.5	1.5	-0.6

	B: Percentiles assuming that all excluded students would score below the 25th percentile but using nonresponse adjustments for absent students					
	25th	50th	75th	90th	95th	
1992	191.1	217.6	239.2	257.3	269.4	
1990	188.3	211.8	233.8	251.6	263.9	
difference	2.8	5.8	5.4	5.7	5.5	

	C: Percentiles assuming that all excluded and absent students would score below the 25th percentile					
	25th	50th	75th	90th	95th	
1992	184.5	214.8	238.1	257.1	266.4	
1990	180.5	208.2	231.2	250.7	260.0	
difference	4.0	6.6	6.9	6.4	6.4	

* standard errors in parentheses

Table 2: Percentiles of Overall Mathematics Proficiency under three treatments of excluded and absent students Grade 8

	A: Percentiles for nonexcluded students with nonresponse adjustment for absent students*					Percent	
	25th	50th	75th	90th	95th	excluded	absent
1992	242.3(1.3)	268.5(1.4)	293.8(0.9)	314.6(1.0)	326.3(1.8)	5.9	10.2
1990	238.7(1.8)	263.6(1.2)	288.0(1.1)	307.1(1.9)	319.2(1.6)	5.3	10.6
difference	3.6	4.9	5.8	7.5	7.1	0.6	-0.4
	B: Percentiles assuming that all excluded students would score below the 25th percentile but using nonresponse adjustments for absent students						
	25th	50th	75th	90th	95th		
1992	237.3	265.6	292.8	312.8	326.3		
1990	233.4	260.9	286.9	305.3	319.2		
difference	3.9	4.7	5.9	7.5	7.1		
	C: Percentiles assuming that all excluded and absent students would score below the 25th percentile						
	25th	50th	75th	90th	95th		
1992	227.8	262.8	291.1	312.1	324.1		
1990	222.8	257.9	285.3	304.7	317.0		
difference	5.0	4.9	5.8	7.4	7.1		

* standard errors in parentheses

Table 3: Percentiles of Overall Mathematics Proficiency under three treatments of excluded and absent students Grade 12

	A: Percentiles for nonexcluded students with nonresponse adjustment for absent students*					Percent	
	25th	50th	75th	90th	95th	excluded	absent
1992	275.0(1.4)	299.9(1.2)	323.2(1.3)	342.6(1.3)	353.6(1.3)	4.1	17.6
1990	269.9(1.3)	295.5(1.5)	319.2(1.4)	339.6(1.6)	350.5(3.1)	3.6	19.3
difference	5.1	4.4	4.0	3.0	3.1	0.5	-1.7

	B: Percentiles assuming that all excluded students would score below the 25th percentile but using nonresponse adjustments for absent students						
	25th	50th	75th	90th	95th		
1992	271.2	298.0	322.1	342.6	353.6		
1990	266.0	293.8	318.3	337.6	350.5		
difference	5.2	4.2	3.8	5.0	3.1		

	C: Percentiles assuming that all excluded and absent students would score below the 25th percentile					
	25th	50th	75th	90th	95th	
1992	252.8	291.4	319.3	340.2	351.7	
1990	242.2	283.8	312.3	335.1	347.7	
difference	10.6	7.6	7.0	5.1	4.0	

* standard errors in parentheses

Table 4: Not reached rates and percent-correct by block and year for the grade 4 national samples

National - Grade 4						
	% NR - Last Item		Avg % NR		Avg. % Correct	
Block:	1992	1990	1992	1990	1992	1990
MD	5	5	1.3	0.9	43	42
ME	25	22	7.4	6.4	43	40
MF	23	26	3.2	3.2	41	37
MH	16	17	3.9	4.5	57	57

Table 5: Not reached rates and percent-correct by block and year for the grade 8 national samples

National - Grade 8						
	% NR - Last Item		Avg % NR		Avg. % Correct	
Block:	1992	1990	1992	1990	1992	1990
MD	10	8	1.6	1.6	58	54
ME	15	15	3.5	3.3	68	65
MF	8	11	1.4	2.1	56	52
MH	29	28	6.0	6.8	49	48

Table 6: Not reached rates and percent-correct by block and year for the grade 12 national samples

National - Grade 12						
	% NR - Last Item		Avg % NR		Avg. % Correct	
Block:	1992	1990	1992	1990	1992	1990
MD	15	11	2.2	1.5	68	65
ME	40	29	6.8	4.9	56	53
MF	27	29	3.5	3.9	56	52
MH	49	45	8.7	8.4	53	51

Table 7: Not reached rates and percent-correct by block and year for the TSA grade 8 samples

Trial State Assessment - Grade 8						
	% NR - Last Item		Avg % NR		Avg. % Correct	
Block:	1992	1990	1992	1990	1992	1990
MD	12	7	2.0	1.1	58	56
ME	20	15	4.2	3.0	68	67
MF	9	8	1.4	1.3	56	53
MH	35	30	7.3	6.1	50	48

APPENDIX P

GRADE 8 WRITING TREND: INVESTIGATION OF THE CHANGES IN MEAN PROFICIENCY BETWEEN 1988 AND 1990, AND BETWEEN 1990 AND 1992

Grade 8 Writing Trend

Investigation of the Changes in Mean Proficiency

Between 1988 and 1990, and Between 1990 and 1992[1]

Jim Carlson & Gene Johnson

Educational Testing Service

May 1994

---

[1] Bruce Kaplan, Phillip Leung, and Steve Wang performed the analyses reported in this paper. Thanks are also due Bruce Kaplan for comments on the paper.

Grade 8 Writing Trend
Investigation of Changes in Mean Proficiency
Between 1988 and 1990, and Between 1990 and 1992

Jim Carlson & Gene Johnson

This document summarizes the investigations ETS has carried out to study the changes in mean proficiency in the writing trend from 1988 to 1990 and from 1990 to 1992. The investigation focussed on the following questions:

- Are the basic item data consistent with the scale scores?
- Are the data from the different years equally reliable, in the interrater sense?
- Are the scorers from different years equivalently stringent?
- Could changes in demographics of the samples explain the changes?
- Do collateral data show the same trend?

Because standard error estimates are not available for many of the reported statistics the focus is primarily descriptive rather than employing hypothesis-testing procedures. The investigation failed to reveal any evidence that the changes are related to anything other than a real changes in grade 8 students' writing proficiency during this four year period.

The remainder of the document describes the investigations and the findings thereof for each of the above questions.

Are the basic item data consistent with the scale scores?

To investigate this question we calculated the mean score attained by examinees administered each of the six grade-eight trend prompts, and the mean proficiency of students who responded to each of these prompts. The mean item scores (writing trend items are scored on a scale of 0 to 4 as described in Chapter 7 and in Mullis, Dossey, Campbell, Gentile, O'Sullivan, & Latham, 1994) for each of the four years in the writing trend are shown in the first part of Table 1. The mean proficiency scores of students responding to each of the prompts (items) are displayed in the second part of this Table.

Table 1. Item Score and Proficiency Means by Prompt by Year

Item	1984	1988	1990	1992
Item Score Means				
Recreational Opportunity	1.58	1.49	1.48	1.63
Food on Frontier	1.88	1.91	1.84	1.92
Dissecting Frogs	1.94	1.98	1.86	2.08
XYZ Company	2.57	2.49	2.47	2.51
Radio Station	2.04	1.91	1.90	2.04
Appleby House	2.37	2.39	2.32	2.51
Mean	2.06	2.03	1.98	2.12
Standardized Mean	0.34	-0.36	-1.36	1.38
Proficiency Means				
Recreational Opportunity	269.4	263.2	257.4	273.6
Food on Frontier	268.5	264.0	257.0	274.6
Dissecting Frogs	267.2	265.0	257.4	276.0
XYZ Company	267.5	265.4	258.1	275.8
Radio Station	265.4	264.0	256.6	274.1
Appleby House	267.1	264.6	258.1	274.4
Mean	267.5	264.4	257.4	274.8
Standardized Mean	0.24	-0.27	-1.36	1.40

Because the standard error estimates of the mean proficiency values reported in the second-last row of Table 1 are available (Mullis et al., 1994, p. 167) the differences between the last three years of the trend were tested for significance using t tests and controlling for familywise error rate using the Bonferroni procedure. The three differences between pairs of means (1992-1990, 1992-1988, 1990-1988) are all significant at the .05 level with a family size of three contrasts. Hence we conclude that there is a drop in mean writing proficiency in grade 8 between 1988 and 1990, followed by an increase between 1990 and 1992. Furthermore the 1992 mean is greater than the 1988 mean.

In order to facilitate comparisons between the scaled proficiency means and the unscaled item means the mean of means was calculated over prompts for each year, and then these four means were standardized by subtracting their overall mean and dividing by their standard deviations. This was done separately for the item score means and the proficiency means. The result is a set of item score means, one for each year of the trend assessment, that is comparable to a set of proficiency score means, also one for each year of the trend assessment. These standardized means, as well as the means before standardization are shown in Table 1.

Figures 1, 2, and 3 display plots of the data in Table 1. Figure 1 shows the trend in individual item scores across the four years and Figure 2 shows the trend in mean proficiency scores. Finally, Figure 3 displays the standardized mean item scores and standardized mean proficiency scores.

As may be seen from the data in Table 1 and Figures 1 to 3, and particularly in Figure 3 and the standardized means in the Table, the trend in the unscaled item score means is nearly identical to that in the scaled proficiency means. The conclusion to be drawn from these data is that the changes in proficiency means during the period of the trend is not a function of the ETS scaling procedures, but is inherent in the data.

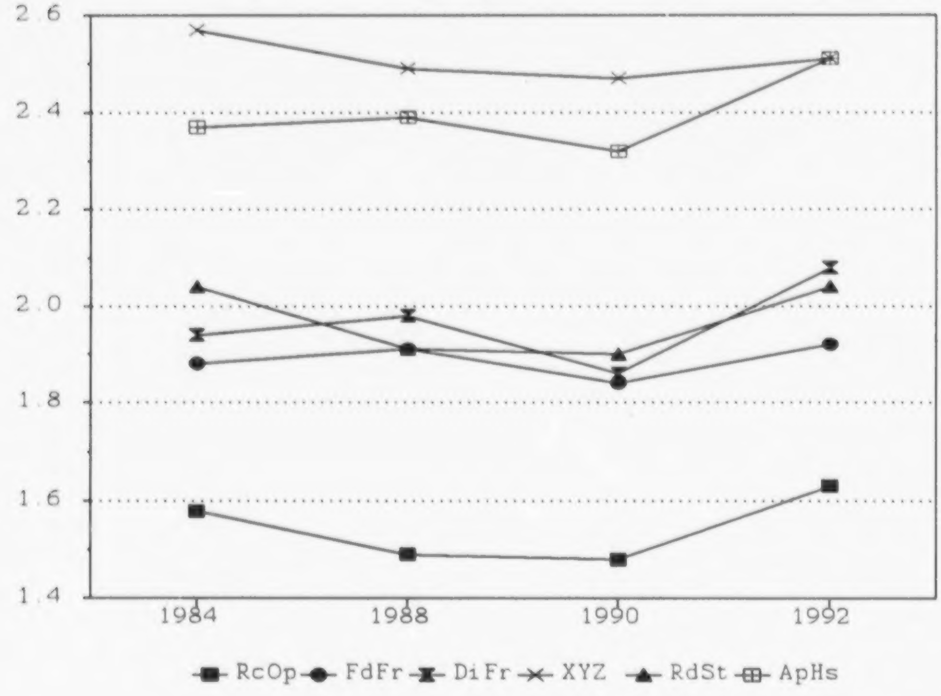


Figure 1

Individual Items: Item Score Means

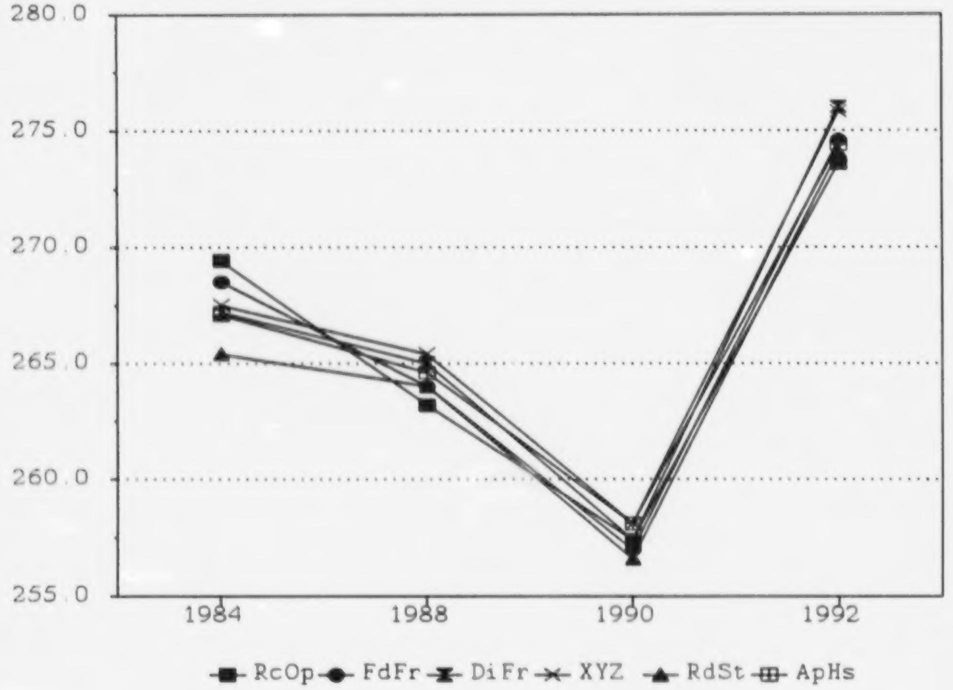


Figure 2

Individual Items: Reporting Proficiency Means

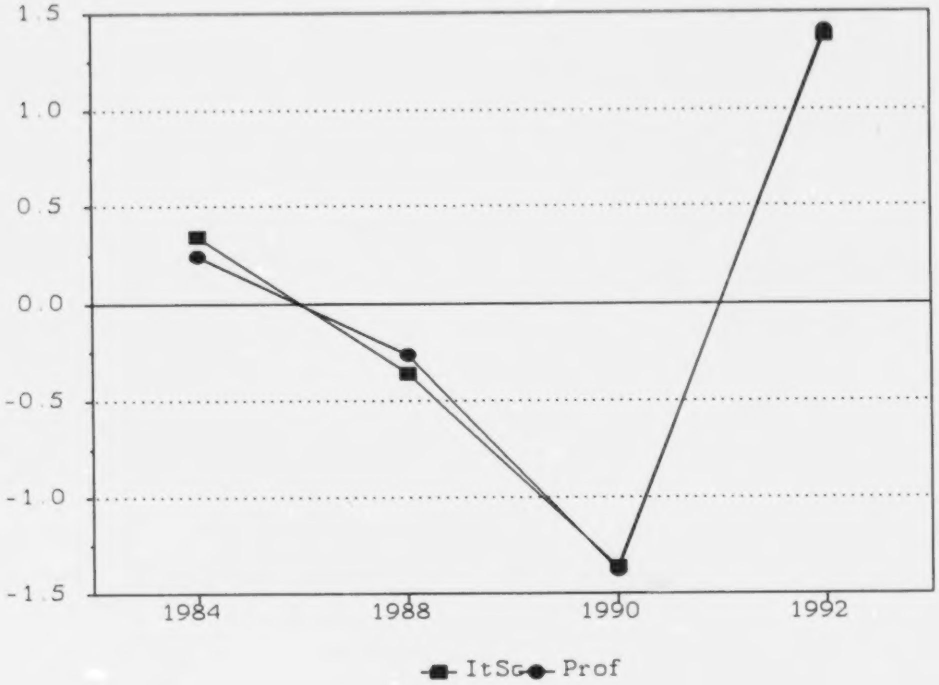


Figure 3

Average Over 6 Items: Standardized Means

Are the data from the different years equally reliable, in the interrater sense?

In order to compare the reliability of the raters' scores for the three assessment years, the 1988 rescore data were examined, in addition to the data from the three years in question. The interrater reliability data result from different scorers rescoring the same papers. For example, a twenty-five percent sample of 1988 papers was rescored by the 1988 scorers, by the 1990 scorers, and again by the 1992 scorers. Similarly, within each year a sample of papers was rescored by different scorers from that year. Table 2 shows the percent agreement between pairs of raters (scorers) and the intraclass correlation coefficient, a measure of interrater reliability.

Although there are some differences in the reliability statistics we do not have standard error estimates for them and thus cannot reach valid conclusions about whether these differences are significant. Some of the statistics are admittedly low, particularly for the Dissecting Frogs prompt. The patterns of scores shown in Figure 1 are very consistent for most of the prompts. Taken together, the information in Figure 1 and Table 2 suggests that if there is any effect of unreliability of scorers it is not strongly related to the proficiency changes between 1988 and 1992. Hence we conclude that the observed increase cannot be attributed to unreliability of the scorers.

Table 2. Interrater Reliability Data

Item	1988 Data by 1988 Scorers		1988 Data by 1990 Scorers		1990 Data by 1990 Scorers		1988 Data by 1992 Scorers		1992 Data by 1992 Scorers	
	Percent Agreement	Reliability	Percent Agreement	Reliability	Percent Agreement	Reliability	Percent Agreement	Reliability	Percent Agreement	Reliability
Recreation Opportunity	85.4	.82	83.0	.81	76.7	.73	79.5	.77	81.9	.82
Food on Frontier	79.9	.68	83.5	.78	72.1	.67	79.4	.68	75.5	.69
Dissecting Frogs	76.1	.64	80.6	.70	66.1	.56	71.2	.54	71.7	.63
XYZ Company	93.5	.92	92.6	.87	86.8	.76	86.2	.76	88.5	.80
Radio Station	87.0	.89	82.0	.79	80.7	.83	85.8	.87	86.0	.89
Appleby House	75.3		75.4	.75	75.9	.72	78.0	.77	82.5	.84
Mean			82.9	.78	76.4	.71	80.0	.73	81.0	.78
Standard Deviation			5.6	.06	7.1	.09	5.6	.11	6.3	.10

Are the scorers from different years equivalently stringent?

To verify that the observed increase is not due to a change in the stringency of the scorers we compared the average scores assigned in the rescoring of the 1988 papers by the 1990 scorers to those of the rescore by the 1992 scorers. We computed, for each item, the difference between the mean score assigned by the 1990 and 1992 scorers on the one hand and the 1988 scorers on the other. These data are presented in Table 3.

Table 3. Mean Differences on 1988 Papers
Comparing 1990 and 1992 Scorers with 1988 Scorers

Prompt	1990-1988 Scorers	1992-1988 Scorers
Recreation Opportunity	-.06	.07
Food on Frontier	.02	-.05
Dissecting Frogs	.00	.06
XYZ Company	.00	-.06
Radio Station	.02	.02
Appleby House	.07	.09
Mean	.01	.02

Examination of the mean differences between 1990 and 1988 scorers, column 2 of Table 3, provides a direct comparison of leniency of the scorers in those two assessment years. As can be seen from the data those differences are quite small and the mean difference over the six prompts is only .01.

Comparison of columns 2 and 3 of Table 3 yields a comparison of the leniency of the 1990 and 1992 assessments' scorers. As can be seen from the data in the table, when the 1990 scorers are compared to the 1992 scorers on grading the same 1988 responses, the differences in mean discrepancy with the original 1988 scorers are very small. Hence we conclude that the 1992 scorers are not being a great deal more lenient than the 1990 scorers, as might be expected if the increase in scores was due to scorer leniency. The differences between the two columns in Table 3 are very small as compared to the differences in 1990 and 1992 mean item scores shown in Table 1.

Could changes in demographics of the samples explain the increase?

To study this question we performed a number of partitioning analyses (Beaton, 1988). These analyses provide a vehicle for studying changes in proportions of examinees in various demographic groups between two assessment years and their relationship to concomitant changes in performance. The partitioning analyses result in estimates of the amount of performance change attributable to each of three sources: population changes, performance change per se (independent of population change), and the interaction of population and performance changes.

A number of different demographic variables were examined by this technique and the results are displayed in Table 4. The first three variables studied are standard NAEP reporting variables. The other variables are from the student background questionnaire and relate specifically to students' perceptions of their writing (I like to write, I'm a good writer, People like what I write), and the amount of writing of various types that the students reported doing during the week previous to administration of the assessment.

Table 4. Partitioning Analysis Results: 1988-1990

Variable Studied	Total Change in Proficiency	Due to Population Changes	Due to Performance Changes	Due to Interaction
Race/Ethnicity	-7.06	-.01	-7.03	-.02
Type of Community	-7.06	-.21	-6.42	-.42
Parents Education	-6.99	-.29	-6.57	-.14
I Like to Write	-7.03	-.19	-6.74	-.09
I'm Good Writer	-6.97	-.04	-6.82	-.11
People Like what I Write	-6.74	-.17	-6.23	-.34
No. Stories Written	-7.33	-.01	-7.32	-.01
No. Essays Written	-7.25	-.07	-7.23	.05
No. Poems Written	-7.38	-.22	-7.10	-.07
No. Plays Written	-7.34	.02	-7.40	.03
No. Letters Written	-7.49	.06	-7.62	.07
No. Book Reports	-7.39	.03	-7.48	.06
No. Other Reports	-7.18	.01	-7.19	.00

Table 5. Partitioning Analysis Results: 1990-1992

Variable Studied	Total Change in Proficiency	Due to Population Changes	Due to Performance Changes	Due to Interaction
Race/Ethnicity	17.43	-.05	17.49	-.01
Type of Community	17.43	.06	17.46	-.09
Parents Education	17.36	1.09	15.95	.31
I Like to Write	16.19	.14	16.03	.02
I'm Good Writer	16.19	.42	16.07	-.30
People Like what I Write	15.82	.64	15.08	.10
No. Stories Written	17.33	-.14	17.39	.08
No. Essays Written	17.28	.01	17.25	.03
No. Poems Written	17.27	-.23	17.37	.13
No. Plays Written	17.14	<.01	17.13	.01
No. Letters Written	17.46	-.31	17.67	.09
No. Book Reports	17.32	-.02	17.33	.01
No. Other Reports	17.07	-.09	17.15	.01

As may be seen from the results reported in Tables 4 and 5, almost no changes in writing proficiency between 1988 and 1990 or between 1990 and 1992 are attributable to changes in demographics of the sample or to interaction of those changes with performance. Rather, the bulk of the change is purely change in performance unassociated with sample demographics. It should be noted in reading these Tables that the total change varies from one partitioning variable to another because not every student responded to every question. The numbers are based on those examinees completing each item of the questionnaire. As in most NAEP analyses weights were used in order that the results best represent the population of concern.

Do collateral data show the same trend?

In addition to the data from the primary trait scoring, analyzed as part of the writing trend report, data are available, for some prompts, on different scoring procedures that were applied to the examinees' writing responses. In particular, some scorers were assigned the task of scoring papers for mechanics, and some to score holistically. If there is a real increase in writing proficiency then it would be expected that this increase would be observable in these other types of scores.

Table 6 displays some results from the mechanics scoring study. This type of scoring was done at the eighth grade on the Recreational Opportunity prompt. Since papers on which examinees responded to this prompt were also scored for primary trait, and holistically, data from these three types of scoring can be compared. The table shows, for each year of the trend assessment in writing, the mean primary trait and holistic scores and, from the mechanics scoring, the mean numbers of sentences and of good sentences. Shown in the second part of the Table are the standardized values of these variables, using the means and standard deviations shown in the two right hand columns of the first part of the Table to perform the standardizations.

Table 6. Mechanics-scoring Sample Data
Recreational Opportunity Prompt

Variable	1984	1988	1990	1992	Mean	Standard Deviation
Item Score Means						
Primary Trait	1.66	1.54	1.51	1.65	1.59	.07
Holistic	3.14	3.42	3.41	3.51	3.37	.14
No. Sentences	4.40	4.90	4.90	5.10	4.83	.26
No. Good Sentences	4.00	4.50	4.40	4.70	4.40	.25
Standardized Means						
Primary Trait	1.06	-.76	-1.21	.91		
Holistic	-1.66	.36	.29	1.01		
No. Sentences	-1.64	.29	.29	1.06		
No. Good Sentences	-1.57	.39	.00	1.18		

A plot of the standardized means is shown in Figure 4. Although not all patterns show a similar trend across all four years of the assessment, the data in the Table and the plot do show a consistent pattern of a decrease between 1988 and 1990 followed by an increase between 1990 and 1992 on all scores. Since standard error estimates are not available for these means no significance tests were performed.

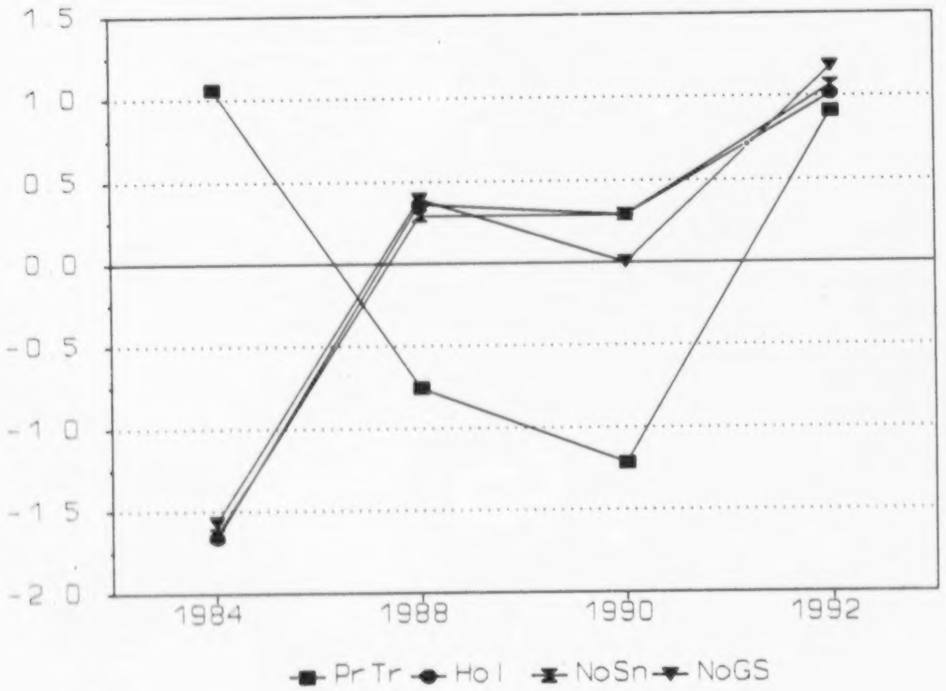


Figure 4

Mechanics Standardized Item Scores: Rec. Opp.

Two of the eighth-grade prompts were scored holistically in addition to the primary trait scoring. This allows us to compare the trend in holistic scores with that of the primary trait scores. Table 6 displays the data for this comparison.

Table 7. Holistic Scoring

Prompt	Item Score Means			
	1984	1988	1990	1992
Recreational Opportunity	3.18	3.37	3.35	3.49
Food on Frontier	3.16	3.22	3.38	3.34

As may be seen in Table 7 and Figure 5 the holistic scoring of the Recreational Opportunity prompt follows the general trend of an increase between 1990 and 1992 whereas the other prompt that was scored holistically, Food on the Frontier, did not follow this trend.

Other data reported in Mullis et al. (1994), especially the data reported in Tables 13.11 and 13.12, show that eighth-grade students are writing various types of assignments once or twice a week significantly more frequently in 1992 than they did in 1984. The types of assignments include: essays, compositions or themes; other reports; letters, stories; and poems. Also, significantly fewer eighth-grade students reported writing no book reports and plays.

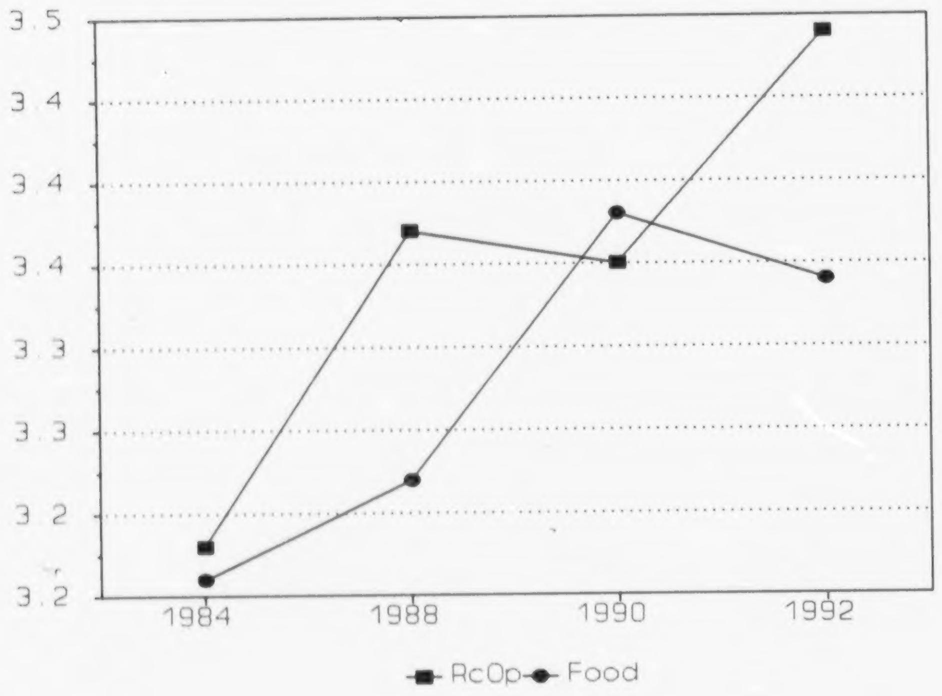


Figure 5

Holistic Scoring: Item Score Means

Conclusions

After examination of data relevant to the five questions posed in the introductory section of this paper, the conclusions are that:

- The changes are not due to scaling procedures because the basic item data show changes that are consistent with the scale scores.
- The changes cannot be attributed to scorer reliability differences because the data from the different years are equally reliable, in the interrater sense.
- The changes cannot be attributed to differential stringency of the raters in the various assessment years.
- Changes in demographics of the samples cannot explain the changes.
- Collateral data tend to show the same trend as the proficiency means.

One possible explanation for the eighth-grade increase between 1990 and 1992, identified in the examination of the collateral data, is that students at this level are writing more than did eighth-grade students in previous years.

REFERENCES CITED IN TEXT

REFERENCES CITED IN TEXT

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika, 52*, 317-332.

Allen, N. L. (1992). Data analysis for the science assessment. In E. G. Johnson & N. L. Allen, *The NAEP 1990 technical report* (No. 21-TR-20). Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress.

Allen, N. L., & Donoghue, J. (1993). Applying the Mantel-Haenszel procedure to complex sample data. Manuscript under review for the *Journal of Educational Measurement*.

Allen, N. L., Mazzeo J., Isham, S. P., Fong, Y. F., & Bowker, D. W. (1993). Data analysis and scaling for the 1992 trial state assessment in reading. In E. G. Johnson, J. Mazzeo, & D. Kline, *Technical report of the NAEP 1992 trial state assessment program in reading*. Washington, DC: National Center for Education Statistics.

Andersen, E. B. (1980). Comparing latent distributions. *Psychometrika, 45*, 121-134.

Anderson, R. C., Hiebert, E. H., Scott, J. A., & Wilkinson, I. A. G. (1984). *Becoming a nation of readers: The report of the commission on reading*. Washington, DC: U. S. Department of Education; The National Institute of Education.

Applebee, A. N., Langer, J. A., Mullis, I. V. S., Latham, A. S., & Gentile, C. A. (1994). *NAEP 1992 writing report card*. Washington, DC: National Center for Education Statistics.

Beaton, A. E. (1988). *The NAEP 1985-86 reading anomaly: A technical report*. Princeton, NJ: Educational Testing Service.

Beaton, A. E. (1987). The NAEP reading scale. In A. E. Beaton, *Implementing the new design: The NAEP 1983-84 technical report* (pp. 381-390). (No. 15-TR-20) Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress.

Beaton, A. E., & Allen, N. L. (1992). Interpreting scales through scale anchoring. *Journal of Educational Statistics, 17*, 191-204.

Beaton, A. E., & Johnson, E. G. (1992). Overview of the scaling methodology used in the National Assessment. *Journal of Educational Measurement, 26(2)*, 163-175.

Beaton, A. E., & Johnson, E. G. (1990). The average response method of scaling. *Journal of Educational Statistics, 15*, 9-38..

Beaton, A. E., & Zwick, R. (1990). *The effect of changes in the National Assessment: Disentangling the NAEP 1985-86 reading anomaly*. (No. 17-TR-21) Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress.

Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: application of an EM algorithm. *Psychometrika, 46*, 443-459.

Bock, R. D., Gibbons, R., & Muraki, E. (1988). Full-information item factor analysis. *Applied Psychological Measurement, 12*, 261-280.

Campbell, J. R., Kapinus, B., & Beatty, A. (1994). *Between the lines: Interviewing children about their literary experiences.* Washington, DC: National Center for Education Statistics.

Carlson, J. E. (1993). *Dimensionality of NAEP instruments that incorporate polytomously-scored items.* Paper presented at the annual meeting of the American Educational Research Association, Atlanta.

Carlson, J. E., & Jirele, T. (1992, April). *Dimensionality of 1990 NAEP mathematics data.* Paper presented at the meeting of the American Educational Research Association, San Francisco.

Cochran, W. G. (1977). *Sampling techniques* (3rd ed.). New York: John Wiley & Sons.

Cohen, J. (1968). Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin, 70*(4), 213-220.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society, Series B, 39*, 1-38.

Dole, J. A., Duffy, G. G., Roehler, L. R., & Pearson, P. D. (1991). Moving from the old to the new: Research on reading comprehension instruction. *Review of Educational Research, 61*(2), 239-264.

Donoghue, J. R., & Mazzeo, J. (1993). *Assessing the properties of longer blocks in the 1992 reading assessment.* Paper presented at the symposium "Analysis of Innovative Data Structures in the National Assessment of Educational Progress" at the annual meeting of the American Educational Research Association, Atlanta, GA.

Educational Testing Service (1987). *ETS standards for quality and fairness.* Princeton, NJ: Author.

Engelen, R. J. H. (1987). *Semiparametric estimation in the Rasch model.* Research Report 87-1. Twente, the Netherlands: Department of Education, University of Twente.

Fraser, C. (1988). *NOHARM A computer program for fitting both unidimensional and multidimensional normal ogive models of latent trait theory.* Armidale, Australia: The University of New England, Center for Behavioral Studies.

Gentile, C. A., & Martin-Rehrmann, J. (1994). *Windows into the classroom: NAEP's 1992 writing portfolio study.* Washington, DC: National Center for Education Statistics.

Guthrie, J. T., & Greaney, V. (1991). Literacy acts. In R. Barr, M. Kamill, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research: Volume II.* New York, NY: Longman.

Hoijtink, H. (1991). *Estimating the parameters of linear models with a latent dependent variable by nonparametric maximum likelihood.* Research Bulletin HB-91-1040-EX. Groningen, The Netherlands: Psychological Institute, University of Groningen.

Holland, P. W. (1985). On the study of differential item performance without IRT. *Proceedings of the Military Testing Association,* October.

Holland, P. W. & Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. In H. Wainer & H. I. Braun (Eds.), *Test validity.* Hillsdale, NJ: Erlbaum.

Johnson, E. G. (1989). Considerations and techniques for the analysis of NAEP data. *Journal of Educational Statistics, 14,* 303-334.

Johnson, E. G. (1990). Data analysis for the writing assessment. In E. G Johnson & R. Zwick (Eds.), *Focusing the new design: the NAEP 1988 technical report* (pp. 267-296). Washington, DC: National Center for Education Statistics, Office of Educational Research and Improvement, U.S. Department of Education.

Johnson, E. G., & Allen, N. L. (1992). *The NAEP 1990 technical report* (No. 21-TR-20). Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress.

Johnson, E. G. & King, B. F. (1987). Generalized variance functions for a complex sample survey. *Journal of Official Statistics, 3,* 235-250.

Johnson, E. G., Mazzeo, J., & Kline, D. L. (1994). *Technical report of the NAEP 1992 trial state assessment program in reading.* Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress.

Johnson, E. G., Mazzeo, J., & Kline, D. L. (1993). *Technical report of the NAEP 1992 trial state assessment program in mathematics.* (No. 23-ST05) Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress.

Johnson, E. G., & Rust, K. F. (1992). Population inferences and variance estimation for NAEP data. *Journal of Educational Statistics.*

Johnson, E. G. & Zwick, R. (1990). *Focusing the new design: The NAEP 1988 technical report.* (No. 19-TR-20) Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress.

Johnson, J. R. (1987). Instrument and item information. In A. E. Beaton, *Implementing the new design: The NAEP 1983-84 technical report* (pp. 119-134). (No. 15-TR-20) Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress.

Kaplan, B. A., Beaton, A. E., Johnson, E. G., & Johnson, J. R. (1988). *National Assessment of Educational Progress 1986 bridge studies.* (No. 17-TR-11) Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress.

Kennel, T., Mohn, L., Reynolds, L., Smrdel, D., Thayer, B., & Moyer, J. (1993). *Report of processing and professional scoring activities, NAEP - 1991-92.* Iowa City, IA: National Computer Systems.

Kish, L. (1965). *Survey sampling.* New York: John Wiley & Sons.

Kish, L. & Frankel, M. R. (1974). Inference from complex samples. *Journal of the Royal Statistical Society, Series B, 36,* 1-22.

Krewski, D. & Rao, J. N. K. (1981). Inference from stratified samples: Properties of linearization, jackknife and balanced repeated replication. *Annals of Statistics, 9,* 1010-1019.

Laird, N. M. (1978). Nonparametric maximum likelihood estimation of a mixing distribution. *Journal of the American Statistical Association, 73,* 805-811.

Lindsey, B., Clogg, C. C., & Grego, J. (1991). Semiparametric estimation in the Rasch model and related exponential response models, including a simple latent class model for item analysis. *Journal of the American Statistical Association, 86,* 96-107.

Little, R. J. A., & Rubin, D. B. (1983). On jointly estimating parameters and missing data. *American Statistician, 37,* 218-220.

Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data.* New York, NY: John Wiley & Sons.

Lord, F. M. (1980). *Applications of item response theory to practical testing problems.* Hillsdale, NJ: Lawrence Erlbaum Associates.

McDonald, R. P. (1962). A general approach to nonlinear factor analysis. *Psychometrika, 27,* 397-415.

McDonald, R. P. (1967). Nonlinear factor analysis. *Psychometric Monographs,* No. 15.

McDonald, R. P. (1981). The dimensionality of tests and items. *British Journal of Mathematical and Statistical Psychology, 34,* 100-117.

McDonald, R. P. (1982a). Linear versus nonlinear models in item response theory. *Applied Psychological Measurement, 6,* 379-396.

McDonald, R. P. (1982b). Unidimensional and multidimensional models for item response theory. In D. J. Weiss (Ed.), *Proceedings of the 1982 Computerized Adaptive Testing Conference* (pp. 127-148). Minneapolis: University of Minnesota, Department of Psychology, Psychometric Methods Program.

McDonald, R. P. (1985). *Factor analysis and related methods.* Hillsdale, NJ: Lawrence Erlbaum.

Messick, S. J., Beaton, A. E., & Lord, F. M. (1983). *NAEP reconsidered: A new design for a new era.* (NAEP Report 83-1) Princeton, NJ: Educational Testing Service.

Mislevy, R. J. (1985). Estimation of latent group effects. *Journal of the American Statistical Association, 80,* 993-997.

Mislevy, R. J. (1990). Scaling procedures. In E.G. Johnson and R. Zwick, *Focusing the new design: The NAEP 1988 technical report* (No. 19-TR-20). Princeton, NJ: National Assessment of Educational Progress, Educational Testing Service.

Mislevy, R. J. (1991). Randomization-based inferences about latent variables from complex samples. *Psychometrika, 56*(2), 177-196.

Mislevy, R. J., Beaton, A. E., Kaplan, B., & Sheehan, K. M. (1992). Estimating population characteristics from sparse matrix samples of item responses. *Journal of Educational Measurement, 29*(2), 133-161.

Mislevy, R. J., & Bock, R. D. (1982). *BILOG: Item analysis and test scoring with binary logistic models* [Computer program]. Mooresville, IN: Scientific Software.

Mislevy, R. J., Johnson, E. G., & Muraki, E. (1992). Scaling procedures in NAEP. *Journal of Educational Statistics, 17(2),* 131-154.

Mislevy, R. J., & Sheehan, K. M. (1987). Marginal estimation procedures. In A. E. Beaton, *Implementing the new design: The NAEP 1983-84 technical report* (pp. 293-360). (No. 15-TR-20) Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress.

Mislevy, R. J., & Wu, P-K. (1988). *Inferring examinee ability when some item responses are missing.* (ETS Research Report RR-88-48-ONR) Princeton, NJ: Educational Testing Service.

Mullis, I. V. S. (1990). *The NAEP guide.* Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress.

Mullis, I. V. S., Dossey, J. A., Campbell, J. R., Gentile, C. A., O'Sullivan, C., & Latham, A. S. (1994). *NAEP 1992 trends in academic progress: Achievement of U.S. students in science, 1969 to 1992; mathematics, 1973 to 1992; reading, 1971 to 1992; writing, 1984 to 1992.* Washington, DC: National Center for Education Statistics.

Mullis, I. V. S., Dossey, J. A., Owen, E. H., & Phillips, G. W. (1993). *NAEP's 1992 mathematics report card for nation and the states.* Washington, DC: National Center for Education Statistics.

Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement, 16(2),* 159-176.

Muraki, E., & Bock, R. D. (1991). *PARSCALE: Parameter scaling of rating data.* Chicago, IL: Scientific Software, Inc.

Muraki, E. & Grima, A M. (1993, May). *Extension of TBLT procedure to the generalized partial credit model.* Paper presented at the NAEP Design and Analysis Committee Meeting, Washington, DC.

Muthén, B. (1991). *Issues in using NAEP mathematics items to study achievement dimensionality, within-grade differences, and across-grade growth.* Report presented to the Design and Analysis Committee for NAEP.

National Academy of Education. (1992). *Assessing student achievement in the states. The first report of the National Academy of Education Panel on the evaluation of the NAEP Trial State Assessment: 1990 Trial State Assessment.* Stanford, CA: Author.

National Assessment of Educational Progress (1987). *Mathematics objectives: 1985-86 assessment.* Princeton, NJ: Educational Testing Service.

National Assessment of Educational Progress (1988). *Mathematics objectives, 1990 assessment.* Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress.

National Assessment of Educational Progress. (1992). *National Assessment of Educational Progress 1992 policy information framework.* Princeton, NJ: National Assessment of Educational Progress, Educational Testing Service.

National Assessment of Educational Progress (1992). *1992 policy information framework.* Princeton, NJ: National Assessment of Educational Progress, Educational Testing Service.

National Assessment Governing Board. (1992a). *Writing framework for the 1992 National Assessment of Educational Progress.* Washington, DC: National Assessment Governing Board, United States Department of Education.

National Assessment Governing Board. (1992b). *Reading framework for the 1992 National Assessment of Educational Progress.* Washington, DC: United States Department of Education.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics.* Reston, VA: Author.

Nelson, J., & Zwick, R. (1989, April). *The Mantel-Haenszel delta difference statistic and its standard error under complex sampling.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Rock, D. A. (1991, November). *Subscale dimensionality*. Paper presented at the meeting of the Design and Analysis Committee of the National Assessment of Educational Progress, Washington, DC.

Rogers, A. M. (1991). *NAEP-MGROUP: Enhanced version of Sheehan's software for the estimation of group effects in multivariate models* [Computer program]. Princeton, NJ: Educational Testing Service.

Rogers, A. M., Barone, J. L., & Kline, D. L. (1990). *A guide to the NAEP item information database*. Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress.

Rogers, A. M., Barone, J. L., & Kline, D. L. (1989). *A guide to the NAEP restricted-use data files*. Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress.

Rogers, A. M., Kline, D. L., Barone, J. L., Mychajlowycz, A. W., & Forer, D. C. (1989). *A primer for the NAEP item information database*. Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress.

Rogers, A. M., Kline, D. L., Johnson, E. G., Mislevy, R. M., & Rust, K. F. (1994). *National Assessment of Educational Progress 1992 secondary-use data files user guide*. Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress.

Rogers, A. M., & Nelson, J. E. (1990). *NAEP-BILOG: Enhanced version of Mislevy and Bock's software for item analysis and test scoring with binary logistic models* [Computer program]. Princeton, NJ: Educational Testing Service.

Rosenbaum P. R. (1984). Testing the conditional independence and monotonicity assumptions of item response theory. *Psychometrika, 49*, 425-436.

Rosenbaum, P. R. (1985). Comparing distributions of item responses for two groups. *British Journal of Mathematical and Statistical Psychology, 38*, 206-215.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: John Wiley & Sons.

Rubin, D. B. (1991). EM and beyond. *Psychometrika, 56*, 241-254.

Rust, K. (1985). Variance estimation for complex estimators in sample surveys. *Journal of Official Statistics, 1(4)*, 381-397.

Rust, K. F., Bethel, J., Burke, J., & Hansen, M. H. (1990). *The 1988 National Assessment of Educational Progress--Sampling and weighting procedures, final report*. Rockville, MD: Westat, Inc.

Rust, K. F., Burke, J., & Fahimi, M. (1992). *1990 National Assessment of Educational Progress sampling and weighting procedures, part 2: National assessment, final report.* Rockville, MD: Westat, Inc.

Satterthwaite, F. E. (1941). Synthesis of variance. *Psychometrika, 6*, 309-316.

Sheehan, K. M. (1985). *MGROUP: Estimation of group effects in multivariate models* [Computer program]. Princeton, NJ: Educational Testing Service.

Stocking, M. L., & Lord, F. M. (1983). Developing a common metric in item response theory. *Applied Psychological Measurement, 7*, 201-210.

Tanner, M., & Wong, W. (1987). The calculation of posterior distributions by data augmentation (with discussion). *Journal of the American Statistical Association, 82*, 528-550.

Thomas, N. (1993). Asymptotic corrections for multivariate posterior moments with factored likelihood functions. *Journal of Computational and Graphical Statistics, 2*, 309-322.

Wallace, L., & Rust, K. F. (1994). *1992 National Assessment of Educational Progress sampling and weighting procedures, part 2: National assessment, final report.* Rockville, MD: Westat, Inc.

Wainer, H., & Kiely, G. (1987). Item clusters and computerized adaptive testing: A case for testlets. *Journal of Educational Measurement, 24*, 185-202.

Wang, M. (1993). *Scaling the 50-minutes Writing Items in the 1992 NAEP Main assessment.* Technical memorandum. Princeton, NJ: Educational Testing Service.

Westat, Inc. (1984). *The private school survey. Final report.* Submitted to the National Center for Education Statistics, U. S. Department of Education. Rockville, MD: Author.

Wilson, D. T., Wood, R., & Gibbons, R. (1991). *TESTFACT: Test scoring, Item Statistics, and Item Factor Analysis 386/486 Version.* Mooresville, IN: Scientific Software.

Wingersky, M., Kaplan, B. A., & Beaton, A. E. (1987). Joint estimation procedures. In A. E. Beaton, *Implementing the new design: The NAEP 1983-84 technical report* (pp. 285-292). (No. 15-TR-20) Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress.

Wolter, K. M. (1985). *Introduction to variance estimation.* New York: John Wiley & Sons.

Yamamoto, K. (1988). Science data analysis. In A. E. Beaton, *Expanding the new design: The NAEP 1985-86 technical report* (pp. 243-255). (No. 17-TR-20) Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress.

Yamamoto, K. (1990). Mathematics and science trend data analysis. In A. E. Beaton & R. Zwick, *The effect of changes in the National Assessment: Disentangling the NAEP 1985-86 reading anomaly* (pp. 111-144). Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress.

Yamamoto, K., & Jenkins, F. (1992). Data analysis for the mathematics assessment. In E. G. Johnson & N. L. Allen, *The NAEP 1990 technical report* (No. 21-TR-20). Washington, DC: National Center for Education Statistics.

Zwick, R. (1986). *Assessment of the dimensionality of NAEP year 15 reading data.* (ETS Research Report No. 86-4). Princeton: Educational Testing Service.

Zwick, R. (1987a). Assessing the dimensionality of NAEP reading data. *Journal of Educational Measurement, 24*, 293-308.

Zwick, R. (1987b). Use of the NAEP almanacs. In A. E. Beaton, *Implementing the new design: The NAEP 1983-84 technical report* (pp. 517-522). (No. 15-TR-20) Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress.

Zwick, R. (1988). *Professionally scored items.* Technical memorandum.

Zwick, R. & Ercikan, K. (1989). Analysis of differential item functioning in the NAEP history assessment. *Journal of Educational Measurement, 26*, 55-66.

Zwick, R. & Grima, A. (1991). *Policy for differential item functioning (DIF) analysis in NAEP.* Technical memorandum.

Zwinderman, A. H. (1991). Logistic regression Rasch models. *Psychometrika, 56*, 589-600.

1130

☆ U.S. GOVERNMENT PRINTING OFFICE: 1994 — 300-807/23104

United States
Department of Education
Washington, DC 20208–5651

Official Business
Penalty for Private Use, $300

NCES 94–490

1131

BEST COPY AVAILABLE